

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 14:57:12 ; Search time 4329.72 Seconds
(without alignments)
16865.305 Million cell updates/sec

Title: US-08-731-499-3
Perfect score: 1507
Sequence: 1 GCAGTTCCTGGATTGACT.....GGTCTTGCACTCCAGGGTCC 1507

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pac.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1504.8	99.9	1507	6	AR070326 Sequence
2	1504.8	99.9	1507	6	BD085727 Genes fro
3	1387.2	92.1	10365	6	BD085733 Genes fro
4	1387.2	92.1	105023	2	AC116668 Trypanoso
5	1387.2	92.1	121143	9	AF312915 Homo sapi
6	1387.2	92.1	128871	9	AL157838 Human DNA
7	1337.2	88.7	5632	6	AR157086 Sequence
8	1337.2	88.7	5632	9	AF041259 Homo sapi
9	1337.2	88.7	5633	6	CQ726295 Sequence
10	1108.6	73.6	3186	6	BD085734 Genes fro
11	614.2	40.8	184420	2	AC118373 Rattus no
12	614.2	40.8	246332	2	AC099079 Rattus no
13	573.6	38.1	114285	10	AL844576 Mouse DNA
14	573.6	38.1	208979	2	AC023610 Mus muscu
15	217.8	14.5	3512	5	BC072191 Xenopus 1
16	217.4	14.4	2750	5	AF419155 Xenopus 1
17	197.2	13.1	298	6	AX895965 Sequence
18	197.2	13.1	298	6	BD031498 Sequence
19	156.4	10.4	469	6	BD060160 Secreted

20	114.4	7.6	212026	2	AC136666	AC136666 Rattus no	
21	114.4	7.6	244851	2	AC094607	AC094607 Rattus no	
C	22	114.4	7.6	268559	2	AC112808	AC112808 Rattus no
23	105.4	7.0	4934	10	BC053104	BC053104 Mus muscu	
C	24	105.4	7.0	175110	10	AC142100	AC142100 Mus muscu
25	104.8	7.0	5982	6	CQ714295	CQ714295 Sequence	
C	26	104.8	7.0	112797	9	AC018413	AC018413 Homo sapi
C	27	104.8	7.0	178717	2	AC024396	AC024396 Homo sapi
28	104.8	7.0	211981	9	AC009716	AC009716 Homo sapi	
29	103.2	6.8	6033	9	D86975	D86975 Homo sapien	
30	97	6.4	8511	5	BC081149	BC081149 xenopus 1	
C	31	79.4	5.3	165094	2	CR547124	CR547124 Danio rer
32	79.4	5.3	174225	5	BX005071	BX005071 Zebrafish	
C	33	79.4	5.3	192578	5	AL732455	AL732455 Zebrafish
C	34	79.4	5.3	243887	2	AC024898	AC024898 Homo sapi
35	78.6	5.2	4265	10	BC033594	BC033594 Mus muscu	
36	78.6	5.2	4417	10	AK122270	AK122270 Mus muscu	
37	78.6	5.2	195922	10	AC140213	AC140213 Mus muscu	
38	74.6	5.0	256281	2	AC094062	AC094062 Rattus no	
C	39	74.6	5.0	266956	2	AC133822	AC133822 Rattus no
C	40	73.2	4.9	171239	2	AC150136	AC150136 Gallus ga
41	73.2	4.9	190305	2	AC150054	AC150054 Gallus ga	
42	70	4.6	70	6	CQ876260	CQ876260 Sequence	
43	70	4.6	4935	6	CQ718121	CQ718121 Sequence	
44	70	4.6	4935	9	AB002388	AB002388 Homo sapi	
45	70	4.6	163120	9	AC011454	AC011454 Homo sapi	

ALIGNMENTS

RESULT 1
AR070326
LOCUS AR070326 1507 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 3 from patent US 5892010.
ACCESSION AR070326
VERSION AR070326.1 GI:7221214
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1507)
AUTHORS Gray,J., Collins,C., Hwang,S.-i., Godfrey,T., Kowbel,D. and
Rommens,J.
TITLE Genes from the 20Q13 amplicon and their uses
JOURNAL Patent: US 5892010-A 3 06-APR-1999;
FEATURES Location/Qualifiers
source
1..1507
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match	99.9%	Score	1504.8	DB	6	Length	1507
Best Local Similarity	100.0%	Pred. No.	0	Mismatches	0	Indels	0
Matches	1507	Conservative	0	0	0	Gaps	0
Qy	1	GCAGTTCCTGGATTGACTTCTTCTCAATTTGAACACTCATTCAATGGACACAAAGAG	60				
Db	1	GCAGTTCCTGGATTGACTTCTTCTCAATTTGAACACTCATTCAATGGACACAAAGAG	60				
Qy	61	CACATAATGCTTTGCTGATTCATATTTGAATTCGAGGCATTTGGGAACCTCTGATGCTTG	120				
Db	61	CACATAATGCTTTGCTGATTCATATTTGAATTCGAGGCATTTGGGAACCTCTGATGCTTG	120				
Qy	121	TTTGTGGAAGAACCACTGACACCACTCACTGAGCTTCTTAAAGTTCGAGAGATTAGAG	180				
Db	121	TTTGTGGAAGAACCACTGACACCACTCACTGAGCTTCTTAAAGTTCGAGAGATTAGAG	180				
Qy	181	GACTATACACTTCTTTTGAACCTTTTATAATAAATTTGCTCTGCTTGTGTAACCCAGG	240				
Db	181	GACTATACACTTCTTTTGAACCTTTTATAATAAATTTGCTCTGCTTGTGTAACCCAGG	240				
Qy	241	ACTGTTAGAGGGTGAGTGACAGGTCTTACAGTGGGCTTAATCAAACCTCAGAAATTTGCC	300				

Db 241 ACTGTTAGGGGTGAGTGACAGGTCTTACAGTGGCTTAATCCAACTCCAGAAATTGGCC 300
Qy 301 AACGGAACCTTTGAGATTATATGCAATCGAAGTGAAGGAAACATGCCAACTCAATCCCT 360
Db 301 AACGGAACCTTTGAGATTATATGCAATCGAAGTGAAGGAAACATGCCAACTCAATCCCT 360
Qy 361 CTTAATGTACATGGATGCCAAGAGTGAATGGCAGCTCTCTTGGCAGTCCGATGGAGATG 420
Db 361 CTTAATGTACATGGATGCCAAGAGTGAATGGCAGCTCTCTTGGCAGTCCGATGGAGATG 420
Qy 421 GAGATGCCCTTGTCAATGAAGGCGCCNCTGTTGTCAATTCGAGCTACACAAAGAAAAA 480
Db 421 GAGATGCCCTTGTCAATGAAGGCGCCNCTGTTGTCAATTCGAGCTACACAAAGAAAAA 480
Qy 481 ATGTCATCCGAATCGAGGGGAATATGCCCTTGGATTGCAATTCGAGCGACGACCTTC 540
Db 481 ATGTCATCCGAATCGAGGGGAATATGCCCTTGGATTGCAATTCGAGCGACGACCTTC 540
Qy 541 ACACATTCAGAGACCTTTAATAAATGCTTTAATGCAACACCGGCTACCTCTGTGAA 600
Db 541 ACACATTCAGAGACCTTTAATAAATGCTTTAATGCAACACCGGCTACCTCTGTGAA 600
Qy 601 CCAGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGGA 660
Db 601 CCAGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGGA 660
Qy 661 ACAGAACCTCCCAAGGAAAGAAATGCAAGGAAATGAATTTAGCTGTGAGGTATGTGG 720
Db 661 ACAGAACCTCCCAAGGAAAGAAATGCAAGGAAATGAATTTAGCTGTGAGGTATGTGG 720
Qy 721 CAGACATTTAGAGTCGCTTTTGTATGTTGAGATCCACATGAGAACACACAAAGATCTTTC 780
Db 721 CAGACATTTAGAGTCGCTTTTGTATGTTGAGATCCACATGAGAACACACAAAGATCTTTC 780
Qy 781 ACTTACGGGTGTAAATGTGCGGAAGAAAGATTCAGGAGCCTTGGTTTCTTAAAAATCAC 840
Db 781 ACTTACGGGTGTAAATGTGCGGAAGAAAGATTCAGGAGCCTTGGTTTCTTAAAAATCAC 840
Qy 841 ATGCGGACRCATATGCGAATCGGGGCGCAGAACAAATGCGAGCAAGGCTTGAGAGT 900
Db 841 ATGCGGACRCATATGCGAATCGGGGCGCAGAACAAATGCGAGCAAGGCTTGAGAGT 900
Qy 901 AGTCAGCAACGATCAACGAGTCTGTCAGGTGCGACGCGCGCAGAGCATCTCTCTCT 960
Db 901 AGTCAGCAACGATCAACGAGTCTGTCAGGTGCGACGCGCGCAGAGCATCTCTCTCT 960
Qy 961 TGC AAAATCTGCATGGTTTGTGGCTTCTCTATTTCCAAAATAAGAAAGTCTAAATTGAGCAC 1020
Db 961 TGC AAAATCTGCATGGTTTGTGGCTTCTCTATTTCCAAAATAAGAAAGTCTAAATTGAGCAC 1020
Qy 1021 CGCAAGGTGCACACAAAAAATCTCTTTCGTACAGCAGCGCGCAGACGACTCTCCA 1080
Db 1021 CGCAAGGTGCACACAAAAAATCTCTTTCGTACAGCAGCGCGCAGACGACTCTCCA 1080
Qy 1081 CAAGGAGGAATGCGCTCTCGAGGGAGGACTTCTGCGAGTGTTCCTCACTTGAGACCAAAA 1140
Db 1081 CAAGGAGGAATGCGCTCTCGAGGGAGGACTTCTGCGAGTGTTCCTCACTTGAGACCAAAA 1140
Qy 1141 TCTCAACCTGAAACGGGGAAGAGCCTGTGAGATGCATCCCTCAGCTCGATCCGTTACC 1200
Db 1141 TCTCAACCTGAAACGGGGAAGAGCCTGTGAGATGCATCCCTCAGCTCGATCCGTTACC 1200
Qy 1201 ACCTTCAGGCTTGCGCAKCTGCTACCAAGAAWAGTTGCCATTTGCCAAGAGTGAAG 1260
Db 1201 ACCTTCAGGCTTGCGCAKCTGCTACCAAGAAWAGTTGCCATTTGCCAAGAGTGAAG 1260
Qy 1261 GAAATGGGGCAAGAGGAGGACCCACACACGAGTTCGAGTTCGAGAGGAGCTTGA 1320
Db 1261 GAAATGGGGCAAGAGGAGGACCCACACGAGTTCGAGTTCGAGAGGAGCTTGA 1320
Qy 1321 GAAACAAATAAGAACCATTTGTGACGGCTCTTCGCAAGAGAAAGAGAGTGC AAAACACTCC 1380

1321 GAAACAAATAAGAACCATTTGTGACGGCTCTTCGCAAGAGAAAGAGAGTGC AAAACACTCC 1380
Qy 1381 CACGGCGAAGCGCCCTCGTGGACCGGATCCCAAGTTACCACTAGCAAGGAGAGCC 1440
Db 1381 CACGGCGAAGCGCCCTCGTGGACCGGATCCCAAGTTACCACTAGCAAGGAGAGCC 1440
Qy 1441 ACTCACTGCTCCGAGTGGGCAAGCTTTTCAAGACCTTACCACCACTGCTGTTGCACTCC 1500
Db 1441 ACTCACTGCTCCGAGTGGGCAAGCTTTTCAAGACCTTACCACCACTGCTGTTGCACTCC 1500
Qy 1501 AGGGTCC 1507
Db 1501 AGGGTCC 1507

RESULT 2
BD085727
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BD085727 1507 bp DNA linear PAT 27-AUG-2002
Genes from the 20q13 amplicon and their uses.
BD085727
BD085727.1 GI:22631337
JP 2001524802-A/3.
synthetic construct
synthetic construct
other sequences; artificial sequences.
1 (bases 1 to 1507)
Gray,J.W., Collins,C.C., Hwang,S.I., Godfrey,T., Kowbel,D. and Rommens,J.
Genes from the 20q13 amplicon and their uses
Patent: JP 2001524802-A 3 04-DEC-2001;
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
OS Artificial Sequence
PN JP 2001524802-A/3
PD 04-DEC-2001
PF 15-JUL-1997 JP 1998506264
PR 15-JUL-1996 US 08/680395,16-OCT-1996 US 08/731499 PR
17-JAN-1997 US 08/785532
PI JOE W GRAY, COLIN CONRAD COLLINS, SOO IN HWANG, TONY GODFREY, PI DAVID KOWBEL,
PI JOHANNA ROMMENS
PC C12N15/11, C12Q1/68, A61K48/00
CC Description of Artificial Sequence: cc49
CC N is A, G, T, or U
FH Key Location/Qualifiers
FT modified base (447).
LOCATION/Qualifiers
1..1507
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

FEATURES
source

ORIGIN

Query Match 99.9%; Score 1504.8; DB 6; Length 1507;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1507; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCAGTCTCGGATTGACTTCTTGTCTCAATTGAAACACTCATTTCAATGGAGACAAAAG 60
Db 1 GCAGTCTCGGATTGACTTCTTGTCTCAATTGAAACACTCATTTCAATGGAGACAAAAG 60
Qy 61 CACTAATCTTTGTGCTGATTCAATTTGAATTCGAGGCACTGGGAACCTGTATGCCCTTG 120
Db 61 CACTAATCTTTGTGCTGATTCAATTTGAATTCGAGGCACTGGGAACCTGTATGCCCTTG 120
Qy 121 TTTGTGGAAGAACACAGTGACACCATCACTGAGCTTCTTAAAGTTTCGAGAGTTAGAG 180
Db 121 TTTGTGGAAGAACACAGTGACACCATCACTGAGCTTCTTAAAGTTTCGAGAGTTAGAG 180
Qy 181 GACTATACACTTCTTTTGAACCTTTTATAATAATTTGCTCTGTTTGGAAACCCAGG 240
Db 181 GACTATACACTTCTTTTGAACCTTTTATAATAATTTGCTCTGTTTGGAAACCCAGG 240
Qy 241 ACTGTTAGAGGGTGAAGTGACAGGTCTTTACAGTGGCCTTTAATCCAACTCCAGAAATTGCC 300

Db 241 ACTGTTAGAGGGTGAAGTGAAGTCTTATCAAGTGGCTTAAATCAATCTCAGAAATTTGCC 300
QY 301 AACCGAACTTTGAGATTATATCAATCGAAAGTGACAGGAAACATGCCAATCTCAATCCCT 360
Db 301 AACCGAACTTTGAGATTATATCAATCGAAAGTGACAGGAAACATGCCAATCTCAATCCCT 360
QY 361 CTTAATGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
Db 361 CTTAATGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
QY 421 GAGATGCTTGTCAATGAAGAGGCGCNCCTGTTGTCAATTCGAGCTACACAAAGAAAAA 480
Db 421 GAGATGCTTGTCAATGAAGAGGCGCNCCTGTTGTCAATTCGAGCTACACAAAGAAAAA 480
QY 481 ATGTCATTCGAATCGAGGGGAATATGCTTGTGATGATGATGATGATGATGATGATGATGATG 540
Db 481 ATGTCATTCGAATCGAGGGGAATATGCTTGTGATGATGATGATGATGATGATGATGATGATG 540
QY 541 ACACATTCAGAGAGCTTTAATAAATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Db 541 ACACATTCAGAGAGCTTTAATAAATGATGATGATGATGATGATGATGATGATGATGATGATG 600
QY 601 CCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGTGATGATGATGATGATGATG 660
Db 601 CCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGTGATGATGATGATGATGATG 660
QY 661 ACAGAACCTCCCAAGGAAAGATGCAAGGAAAGATGATGATGATGATGATGATGATGATGATG 720
Db 661 ACAGAACCTCCCAAGGAAAGATGCAAGGAAAGATGATGATGATGATGATGATGATGATGATG 720
QY 721 CAGACATTTAGAGTGCCTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Db 721 CAGACATTTAGAGTGCCTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 780
QY 781 ACTTACGGGTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Db 781 ACTTACGGGTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
QY 841 ATGGGACRCATATGCGCAATCGGGGCGCAAGCAAAATGCAAGCAAGGCTTGAGAGT 900
Db 841 ATGGGACRCATATGCGCAATCGGGGCGCAAGCAAAATGCAAGCAAGGCTTGAGAGT 900
QY 901 AGTCAGCAACATCAACGAGTCTGTCAGGTGTCAGCGGCGGAGAGCATCTCTCTCTCT 960
Db 901 AGTCAGCAACATCAACGAGTCTGTCAGGTGTCAGCGGCGGAGAGCATCTCTCTCTCTCT 960
QY 961 TGCATAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
Db 961 TGCATAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
QY 1021 CGCAGGTGCACCAAAATGCTTTTGGTACCGAGCGCGCAGAGCATCTCTCTCTCTCT 1080
Db 1021 CGCAGGTGCACCAAAATGCTTTTGGTACCGAGCGCGCAGAGCATCTCTCTCTCTCTCT 1080
QY 1081 CAAGGAGGAATGCCCTCTCGAGGAGGACTTCTGAGTGTGATGATGATGATGATGATGATG 1140
Db 1081 CAAGGAGGAATGCCCTCTCGAGGAGGACTTCTGAGTGTGATGATGATGATGATGATGATG 1140
QY 1141 TCTCACTTGAACCGGGGAAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1200
Db 1141 TCTCACTTGAACCGGGGAAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1200
QY 1201 ACCTTCAGGCTTGGCACTGCTACCAAGAAAGTGGCATTTGGCATTTGGCAAGAGTGAAG 1260
Db 1201 ACCTTCAGGCTTGGCACTGCTACCAAGAAAGTGGCATTTGGCATTTGGCAAGAGTGAAG 1260
QY 1261 GAAATTTGGGCGAAGAGGGAGCAACGACGATTCGAGTTCGAGAGGAGGCTTGGA 1320
Db 1261 GAAATTTGGGCGAAGAGGGAGCAACGACGATTCGAGTTCGAGTTCGAGAGGAGGCTTGGA 1320
QY 1321 GAAACAAATAGAACCATTTGTGAGGCTCTCGCAAGAGAAAGAGAGTGCACATCTCC 1380

Db 1321 GAAACAAATAGAACCATTTGTGAGGCTCTCGCAAGAGAAAGAGTGCACATCTCC 1380
QY 1381 CACGGCGAAGCGCCCTCGTGGACCGGATCCCAAGTTACCCAGTAGCAAGAGAGCC 1440
Db 1381 CACGGCGAAGCGCCCTCGTGGACCGGATCCCAAGTTACCCAGTAGCAAGAGAGCC 1440
QY 1441 ACTCACTCTCGAGTGGCGAAAGCTTTTGAACCTTACACCACTACCACTGGTCTTTGCACTCC 1500
Db 1441 ACTCACTCTCGAGTGGCGAAAGCTTTTGAACCTTACACCACTACCACTGGTCTTTGCACTCC 1500
QY 1501 AGGGTCC 1507
Db 1501 AGGGTCC 1507
RESULT 3
BD085733 10365 bp DNA linear PAT 27-AUG-2002
LOCUS Genes from the 20ql3 amplicon and their uses.
DEFINITION BD085733
ACCESSION BD085733
VERSION BD085733.1 GI:22631343
KEYWORDS JP 2001524802-A/9.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 10365)
AUTHORS Gray, J.W., Collins, C., Hwang, S.I., Godfrey, T., Kowbel, D. and Rommens, J.
TITLE Genes from the 20ql3 amplicon and their uses
JOURNAL Patent: JP 2001524802-A 9 04-DEC-2001;
COMMENT THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
OS Artificial Sequence
PN JP 2001524802-A/9
PD 04-DEC-2001
PR 15-JUL-1997 JP 1998506264
PR 15-JUL-1996 US 08/680395,16-OCT-1996 US 08/731499 PR
17-JAN-1997 US 08/785532
PI JOE W GRAY, COLIN CONRAD COLLINS, SOO IN HWANG, TONY GODFREY, PI
DAVID KOWBEL,
PI JOHANNA ROMMENS
PC C12N15/11, C12Q1/68, A61K48/00
CC Description of Artificial Sequence: Genomic Sequence encoding
CC ZABC1
FH Key Location/Qualifiers
FT source 1..10365
FT source 1..10365 /organism='Artificial Sequence'.
FEATURES
source Location/Qualifiers
1..10365
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'
ORIGIN
Query Match 92.1%; Score 1387.2; DB 6; Length 10365;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1470; Conservative 3; Mismatches 32; Indels 6; Gaps 6;
QY 2 CAGGTTGCTGGATTGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAGAGC 61
Db 2424 CAGGTTGCTGGATTGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAGAGA 2483
QY 62 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAAACCCCTGTATGCTTGT 121
Db 2484 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAAACCCCTGTATGCTTGT 2543
QY 122 TTGTGAAAGAACCGAGTGACACCATCAGCTTCCTTAAAGATTCGAAGAGTTAGAGG 181
Db 2544 TTGTGAAAGAACCGAGTGACACCATCAGCTTCCTTAAAGATTCGAAGAGTTAGAGG 2603
QY 182 ACTATACACTTCTTTTGAACCTTTTATATAAATATTTGCTCTGG- TTTTGGAAACCCAGG 240
Db 2604 ACTATACACTTCTTTTGAACCTTTTATATAAATATTTGCTCTGGTTT TTTGGAACCCAGG 2663

241 ACTGTTAGA-GGGTCAGTGAACAGGCTCTTAC-AGTGGCCCTTAATCAACCTCCAGAAATGTC 298
 Db GCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCCTTATTCCAACTCCAGAAATGTC 2723
 299 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 358
 Db CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 2783
 359 CTCCTTAATGTACATGGAT-GGCCAAGAGTGAATGGCAGCTCTCTTCCAGTCCGATGGAG 417
 Db CTCCTTAATGTACATGGATGGCCCAAGAGTGAATGGCAGCTCTCTTCCAGTCCGATGGAG 2843
 418 ATGGA-GATGCTCTGTCAATGAAGGGCCNCTGTGTGCAATTCGAGCTACACAAAGAA 476
 Db ATGAGGATGCTGTCAATGAAGGGACCGCTGTGTTCATTCGAGCTACACAAAGAA 2903
 477 AAAAATGTCGAATCGAAGGGAATATCCCTTTGGATTCGATGTCATGTCAGCCAGAC 536
 Db AAAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 2962
 537 CTTTACACATTCAGAGACCTTAATAACATGCTTAAATGCAACACCGGCTTACCTCTG 596
 Db CTTTACACATTCAGAGACCTTAATAACATGCTTAAATGCAACACCGGCTTACCTCTG 3022
 597 TGAACCAAGCAGTCTTTCGGGTGAAGCAGAGTATCTCAGTCCGCTTGTATAAAGTCAAGT 656
 Db TGAACCAAGCAGTCTTTCGGGTGAAGCAGAGTATCTCAGTCCGCTTGTATAAAGTCAAGT 3082
 657 GCGAACAGAACTCCCAAGGAAAAGAAATGCAAGAAAATGAATTTAGCTGTGAGGTATG 716
 Db GCGAACAGAACTCCCAAGGAAAAGAAATGCAAGAAAATGAATTTAGCTGTGAGGTATG 3142
 717 TGGCAGACATTTAGATCGCTTTTGTATGTCAGATCCACATGAGACACACAAAGATTC 776
 Db TGGCAGACATTTAGATCGCTTTTGTATGTCAGATCCACATGAGACACACAAAGATTC 3202
 777 TTTTCACTTACGGGTGTAACTGTGCGAAGAAAGATTCAGAGAGCTTGGTTTCTTAAAAA 836
 Db TTTTCACTTACGGGTGTAACTGTGCGAAGAAAGATTCAGAGAGCTTGGTTTCTTAAAAA 3262
 837 TCATATGCGGACRCAATATGCAAAATCGGGGGCAGAAAGCAAACTGCAGCAAGGCTTGA 896
 Db TCATATGCGGACACATAATGCAAAATCGGGGGCAGAAAGCAAACTGCAGCAAGGCTTGA 3322
 897 GAGTAGTCCAGCACAGATCAACAGAGTGTCTCAGGTGCACGCGCGCAGAGCATCTCTC 956
 Db GAGTAGTCCAGCACAGATCAACAGAGTGTCTCAGGTGCACGCGCGCAGAGCATCTCTC 3382
 957 TCCTTGCAAAATCTGCATGTTTGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAAATTGA 1016
 Db TCCTTGCAAAATCTGCATGTTTGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAAATTGA 3442
 1017 GCACCGAGGTGCACACCAAAAAAATGCTTTTGGTACAGACAGCGCGCAGAGCATC 1076
 Db GCACCGAGGTGCACACCAAAAAAATGCTTTTGGTACAGACAGCGCGCAGAGCATC 3502
 1077 TCCACAGGAGGAATGCGCTCCTCGAGGAGGACTTCTCTGAGTGTGTTCAACTTGGAGAC 1136
 Db TCCACAGGAGGAATGCGCTCCTCGAGGAGGACTTCTCTGAGTGTGTTCAACTTGGAGAC 3562
 1137 AAAATCTCACCTGAAACGGGGAAGAGCCGTGTGATGTCATCTCCTCAGCTCGATCCGTT 1196
 Db AAAATCTCACCTGAAACGGGGAAGAGCCGTGTGATGTCATCTCCTCAGCTCGATCCGTT 3622
 1197 CACCACTTCCAGGCTTGGCAKCTGGCTACCAAGAGAAAGTGTGCAATTTGCCAAGAGT 1256
 Db CACCACTTCCAGGCTTGGCAKCTGGCTACCAAGAGAAAGTGTGCAATTTGCCAAGAGT 3682
 1257 GAAGGAATTTGGGCAAGAGGAGCACCACAAACGAGATTCGAGTTCGAGAGAGGAGCT 1316
 Db GAAGGAATTTGGGCAAGAGGAGCACCACAAACGAGATTCGAGTTCGAGAGAGGAGCT 3742
 1317 TGGAGAAACAAATAAGAACCAATTTGTGCGAGGCTCTCTCGCAAGAGAAAGAGTGCAACA 1376

3743 TGGAGAAACAAATAAGGCGAGTGTGTCAGGCTCTCGCAAGAGAAAGAGAGTGCAACA 3802
 QY CTCCACGGCGAAGCGCCCTCCGTGGAGCGGATCCCAAGTATCCAGTAGCAAGAGAA 1436
 Db CTCCACGGCGAAGCGCCCTCCGTGGAGCGGATCCCAAGTATCCAGTAGCAAGAGAA 3862
 1437 GCCCACTCACTGCTCCGAGTGCGGCAAGCTTTTCCAGAACCTTACCACAGCTGCTTCTGCA 1496
 Db GCCCACTCACTGCTCCGAGTGCGGCAAGCTTTTCCAGAACCTTACCACAGCTGCTTCTGCA 3922
 1497 CTCCAGGCTCC 1507
 Db CTCCAGGCTCC 3933

AC116668 105023 bp DNA linear HTG 01-MAY-2002
 Trypanosoma brucei chromosome V clone RPC193-45E22, *** SEQUENCING
 IN PROGRESS ***, 3 unordered pieces.
 AC116668
 VERSION AC116668.5 GI:20376999
 KEYWORDS HTG; HTGS PHASE1.
 SOURCE Trypanosoma brucei
 ORGANISM Trypanosoma brucei
 Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
 Trypanosoma.

REFERENCE
 1 (bases 1 to 105023)
 El-Sayed,N.M., Ghedin,E., Song,J., Larkin,C., Wanless,D., Jones,K.,
 Peterson,J., Hou,L., Zhao,H., Mason,T., Militscher,J., Pai,G., Van
 Aken,S., Uterback,T., Khalak,H.G., Gerard,C., Leech,V., Ullu,E.,
 Melville,S., White,O., Adams,M.D., Donelson,J.E. and Fraser,C.M.
 Trypanosoma brucei GUTat10.1 RPC193-45E22 BAC genomic sequence
 Unpublished
 2 (bases 1 to 105023)
 El-Sayed,N.M., Khalak,H. and Adams,M.D.
 Direct Submission
 Submitted (02-APR-2002) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 3 (bases 1 to 105023)
 El-Sayed,N.M., Khalak,H. and Adams,M.D.
 Direct Submission
 Submitted (01-MAY-2002) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 On May 1, 2002 this sequence version replaced gi:20340472.
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 85781: contig of 85781 bp in length
 * 85782 85806: gap of unknown length
 * 85807 102956: contig of 17150 bp in length
 * 102957 102981: gap of unknown length
 * 102982 105023: contig of 2042 bp in length.

FEATURES
 Location/Qualifiers
 1..105023
 /organism="Trypanosoma brucei"
 /mol_type="genomic DNA"
 /isolate="GUTat10.1"
 /db_xref="taxon:5691"
 /chromosome="V"
 /clone="RPC193-45E22"

ORIGIN
 Query Match 92.1%; Score 1387.2; DB 2; Length 105023;
 Best Local Similarity 97.3%; Pred. No. 0;
 Matches 1470; Conservative 3; Mismatches 32; Indels 6; Gaps 6;

QY 2 CAGGTTGCTGGGATTCGACTTCTTGCTCAATGAAACACTCATTTCAATGGAGACAAAGAGC 61
DB 6708 CAGGTTGCTGGGATTCGACTTCTTGCTCAATGAAACACTCATTTCAATGGAGACAAAGAGC 6767
QY 62 ACTAATGCTTTTGCTGATTCATATTTGAATCGAGGCATTTGGAAACCTGTATGCCCTTCT 121
DB 6768 ACTAATGCTTTTGCTGATTCATATTTGAATCGAGGCATTTGGAAACCTGTATGCCCTTCT 6827
QY 122 TTGTGGAAGAAACCACTGACACCATCACTGAGCTTCTTAAAGTTTCGAAGAAGTTAGAGG 181
DB 6828 TTGTGGAAGAAACCACTGACACCATCACTGAGCTTCTTAAAGTTTCGAAGAAGTTAGAGG 6887
QY 182 ACTATACACTTTCTTTTGAACCTTTTATATAAATATTTGCTCTGG-TTTTGGAACCCAGG 240
DB 6888 ACTATACACTTTCTTTTGAACCTTTTATATAAATATTTGCTCTGGTTTGTGGAACCCAGG 6947
QY 241 ACTGTTTGA-GGGTGAGTGACAGCTCTTAC-AGTGGCCTTTAATCCAACTCCAGAAATTCG 298
DB 6948 GCTGTTAGAGGGTGAGTGACAAAGTCTTACAGTGGCCTTTATTCAACTCCAGAAATTCG 7007
QY 299 CCAACCGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 358
DB 7008 CCAACCGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 7067
QY 359 CTCCTTAATGTACATGAT-GGCCAAGAGTGATTTGGCAGCTCTCTTGCAGTCCGATGGAG 417
DB 7068 CTCCTTAATGTACATGATGGCCAGAAAGTGATTTGGCAGCTCTCTTGGCAGTCCGATGGAG 7127
QY 418 ATGGA-GATGCTCTGTCAATGAAGGGCCNCTGTTGTCAATTCGAGCTACACAAGAA 476
DB 7128 ATGGAAGATGCTCTGTCAATGAAGGGCCGCTGTTGTTCATTCGAGCTACACAAGAA 7187
QY 477 AAAAAATGTCAATCCGAATCAGGGGAATATGCCCTTGGATTGCTCTGAGCCAGAC 536
DB 7188 AAAAAATGTC-ATCCAATCGAGGGGTATATGCCCTTGGATTGCTCTGAGCCAGAC 7246
QY 537 CTTTACACATTCAGAAAGACTTAATAAACAATGCTTTAATGCAACACCGGCCCTACCTCTG 596
DB 7247 CTTTACACATTCAGAAAGACTTAATAAACAATGCTTTAATGCAACACCGGCCCTACCTCTG 7306
QY 597 TGAACACAGATCTCTCGGTTGAACGAGATATCTCAGTCCGCTTGATAAAGTCAAGT 556
DB 7307 TGAACACAGATCTCTCGGTTGAACGAGATATCTCAGTCCGCTTGATAAAGTCAAGT 7366
QY 657 GCGAACAGAACTCCCAAGGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATG 716
DB 7367 GCGAACAGAACTCCCAAGGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATG 7426
QY 717 TGGGCAGACATTTAGAGTCGCTTTTGATTTGATGATCCCATGAGAACACACAAAGATTC 776
DB 7427 TGGGCAGACATTTAGAGTCGCTTTTGATTTGATGATCCCATGAGAACACACAAAGATTC 7486
QY 777 TTTCACTTACGGGTGTAAATGTCGGAAGAGATTCAGAGGACCTTGTCTTTAAAAA 836
DB 7487 TTTCACTTACGGGTGTAAATGTCGGAAGAGATTCAGAGGACCTTGTCTTTAAAAA 7546
QY 837 TCACATGCGGACRATAAATGCGAAATTCGGGGCCAGAAGCAACTGCAGAGGCTTTGA 896
DB 7547 TCACATGCGGACRATAAATGCGGAATTCGGGGCCAGAAGCAACTGCAGAGGCTTTGA 7606
QY 897 GAGTAGTCAGCAACCATCAACGAGTCTGTCAGGTCGTCAGGGCCGAGAGCATCTCTTC 956
DB 7607 GAGTAGTCAGCAACCATCAACGAGTCTGTCAGGTCGTCAGGGCCGAGAGCATCTCTTC 7666
QY 957 TCGTTGCAAAATCTGATGTTGCTGCTTCTTATTTCCAAATAAAGAAGTCTAATGA 1016
DB 7667 TCGTTGCAAAATCTGATGTTGCTGCTTCTTATTTCCAAATAAAGAAGTCTAATGA 7726
QY 1017 GCACCGCAAGGTGCACACCAAAAAAATCTGTTTCGTTACAGCAGCGCGCAGACAGACTC 1076
DB 7727 GCACCGCAAGGTGCACACCAAAAAAATCTGTTTCGTTACAGCAGCGCGCAGACAGACTC 7786
QY 1077 TCCACAAGGAGGAATCCGCTCTCGAGGGAGGACTTCTCTGCAATTTGTTCAACTTGAGACC 1136

DB 7787 TCCACAAGGAGGAATGCCGCTCTCGAGGGAGGACTTCTCTGCACTTGTTCCAACTTGAGACC 7846
QY 1137 AAAATCTCACTCCCTGAAACGGGGAAGAGCTGTGCAGATGCATCCCTCAGCTCGATCCGTT 1196
DB 7847 AAAATCTCACTCCCTGAAACGGGGAAGAGCTGTGCAGATGCATCCCTCAGCTCGATCCGTT 7906
QY 1197 CACCACTTCCAGGCTTGGCACTGTGCTGTACCAAGAAAGTGTGCATTTGCCAAGAAGT 1256
DB 7907 CACCACTTCCAGGCTTGGCAGCTGGCTACCAAGAAAGTGTGCATTTGCCAAGAAGT 7966
QY 1257 GAAGGATTTGGGGCAAGAGGAGCAGCAGCAACGAGATTCGAGTTCCGAGAGGAGCT 1316
DB 7967 GAAGGATTTGGGGCAAGAGGAGCAGCAGCAACGAGATTCGAGTTCCGAGAGGAGCT 8026
QY 1317 TCGAGAAACAAATTAAGAACCATTTGTGAGGCTCTCTGCAAGAGAAAGAGTGCAGAA 1376
DB 8027 TCGAGAAACAAATTAAGGAGCTTTGTGAGGCTCTCTGCAAGAGAAAGAGTGCAGAA 8086
QY 1377 CTCCACGCGGAGCGCCCTCCGTGGAACGCGATCCCAAGTTACCCAGTAGCAAGAGAA 1436
DB 8087 CTCCACGCGGAGCGCCCTCCGTGGAACGCGATCCCAAGTTACCCAGTAGCAAGAGAA 8146
QY 1437 GCCCACTACTGCTCGAGTGGCGCAAGCTTTTCAGAACTTACCACAGCTGCTCTTGA 1496
DB 8147 GCCCACTACTGCTCGAGTGGCGCAAGCTTTTCAGAACTTACCACAGCTGCTCTTGA 8206
QY 1497 CTCCAGGGTCC 1507
DB 8207 CTCCAGGGTCC 8217

RESULT 5
AP312915/c
LOCUS AP312915 121143 bp DNA linear PRI 13-JUN-2001
DEFINITION Homo sapiens chromosome 20 clones 97 and 127, complete sequence.
ACCESSION AP312915
VERSION AP312915.1 GI:11094030
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 121143)
AUTHORS Collins,C., Volik,S., Kowbel,D., Ginzinger,D., Ylstra,B.,
Cloutier,T., Hawkins,T., Predki,P., Martin,C., Wernick,M.,
Kuo,W.L., Alberts,A. and Gray,J.W.
Comprehensive genome sequence analysis of a breast cancer amplicon
Genome Res. 11 (6), 1034-1042 (2001)
21275464
PUBMED 11381030
REFERENCE 2 (bases 1 to 121143)
AUTHORS Volik,S., Collins,C., Gray,J., Wernick,M., Kowbel,D., Stultz,K. and
Martin,C.
Direct Submission
Submitted (10-OCT-2000) Cancer Genetics, UCSF Cancer Center, 2340
Sutter St., Rm. S151, San Francisco, CA 94706, USA
FEATURES
source
1..121143
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="20"
/map="20q13.2"
/clone="97; 127; LBNL h117 and h79"
/note="Isolated in Collins lab and sequenced by the Genome
Center at LBNL"
complement(1..66093)
/note="overlaps Homo sapiens chromosome 20 clone
RP5-823G15 deposited in GenBank Accession Number
AL354993.19"
194..468
repeat_region
/rpt_family="Alu"

COMMENT

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, ENBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/c_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr20>

This sequence is the entire insert of clone RP4-724E16 The true right end of clone RP5-823G15 is at 19684 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP4-724E16 is from the library RPCI-4 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2.

FEATURES

source	Location/Qualifiers
repeat_region	1..54
repeat_region	/note="MLT2A repeat: matches 362. .415 of consensus"
repeat_region	55..368
repeat_region	/note="AluXs repeat: matches 1. .312 of consensus"
repeat_region	369..559
repeat_region	/note="AluSq repeat: matches 126. .312 of consensus"
repeat_region	1137..1270
repeat_region	/note="FLAM C repeat: matches 1. .133 of consensus"
misc_feature	Complement(1396..1840)
repeat_region	1415..1709
repeat_region	/note="AluJo repeat: matches 14. .288 of consensus"
repeat_region	1744..1902
repeat_region	/note="MIR repeat: matches 14. .175 of consensus"
misc_feature	1871..2326
repeat_region	/note="match: GSS: Em:A0673688"
repeat_region	1948..2041
repeat_region	/note="MERSA repeat: matches 6. .101 of consensus"
repeat_region	2622..2787
repeat_region	/note="MERSA repeat: matches 4. .189 of consensus"
repeat_region	2806..2870
repeat_region	/note="MERSA repeat: matches 10. .75 of consensus"
repeat_region	2872..3026
repeat_region	/note="LiMA4 repeat: matches 6017. .6162 of consensus"
repeat_region	3027..3307
repeat_region	/note="AluJo repeat: matches 1. .290 of consensus"
repeat_region	3308..3436
repeat_region	/note="LiMA4 repeat: matches 6162. .6293 of consensus"
repeat_region	3455..3551
repeat_region	/note="MERSA repeat: matches 61. .183 of consensus"
repeat_region	3564..3697
repeat_region	/note="MERSB repeat: matches 14. .146 of consensus"
repeat_region	3871..4004
repeat_region	/note="AluSq/x repeat: matches 1. .134 of consensus"
repeat_region	4249..4274
repeat_region	/note="13 copies 2 mer gt 96% conserved"
repeat_region	4810..5101
repeat_region	/note="AluXs repeat: matches 1. .297 of consensus"
repeat_region	5110..5413
repeat_region	/note="AluSc repeat: matches 1. .305 of consensus"
repeat_region	5492..5648
repeat_region	/note="L2 repeat: matches 2589. .2745 of consensus"
repeat_region	5659..6135
repeat_region	/note="MERIA repeat: matches 1. .488 of consensus"
repeat_region	6164..6225
repeat_region	/note="31 copies 2 mer aa 71% conserved"
repeat_region	6245..6379
repeat_region	/note="AluSg/x repeat: matches 162. .296 of consensus"
repeat_region	6380..6537
repeat_region	/note="79 copies 2 mer aa 60% conserved"
repeat_region	6558..6967
repeat_region	/note="AluSg repeat: matches 1. .309 of consensus"
repeat_region	7320..7542
repeat_region	/note="MERSB repeat: matches 117. .341 of consensus"
repeat_region	7545..7838
repeat_region	/note="AluXs repeat: matches 1. .293 of consensus"
misc_feature	8465..9088
misc_feature	/note="match: GSS: Em:AQ386102"
repeat_region	8523..8814
repeat_region	/note="AluJb repeat: matches 1. .294 of consensus"
repeat_region	9262..9561
repeat_region	/note="AluSq repeat: matches 1. .301 of consensus"
repeat_region	9562..9599
repeat_region	/note="19 copies 2 mer tt 81% conserved"
repeat_region	9606..9733
repeat_region	/note="MER33 repeat: matches 17. .148 of consensus"
repeat_region	9761..9915
repeat_region	/note="MERSA repeat: matches 4. .183 of consensus"
repeat_region	9980..10108
repeat_region	/note="L2 repeat: matches 2622. .2747 of consensus"
repeat_region	10128..10437
repeat_region	/note="AluXs repeat: matches 2. .312 of consensus"
repeat_region	10654..10724
repeat_region	/note="MIR repeat: matches 63. .137 of consensus"
repeat_region	10781..11083
repeat_region	/note="AluXs repeat: matches 1. .306 of consensus"
misc_feature	complement(10897..11319)
misc_feature	/note="match: GSS: Em:AQ499673"
misc_feature	complement(10991..11319)
misc_feature	/note="match: GSS: Em:AQ283579"
repeat_region	11462..11759
repeat_region	/note="AluSq repeat: matches 1. .299 of consensus"
repeat_region	11878..12182
repeat_region	/note="AluXs repeat: matches 1. .300 of consensus"
repeat_region	12183..12587
repeat_region	/note="L1P8 repeat: matches 5757. .6163 of consensus"
repeat_region	13333..13637
repeat_region	/note="MLTIG repeat: matches 3. .338 of consensus"
repeat_region	14201..14507
repeat_region	/note="AluY repeat: matches 1. .308 of consensus"
repeat_region	14540..14694
repeat_region	/note="L2 repeat: matches 2550. .2709 of consensus"
repeat_region	14764..14868
repeat_region	/note="MIR repeat: matches 103. .208 of consensus"
repeat_region	14891..15152
repeat_region	/note="AluJo repeat: matches 12. .299 of consensus"
repeat_region	15162..15293
repeat_region	/note="L2 repeat: matches 1697. .1824 of consensus"
repeat_region	15317..15544
repeat_region	/note="L1MB3 repeat: matches 5943. .6173 of consensus"
repeat_region	16306..16603
repeat_region	/note="AluSc repeat: matches 1. .306 of consensus"
repeat_region	16610..16901
repeat_region	/note="AluXs repeat: matches 1. .292 of consensus"
repeat_region	16989..17035
repeat_region	/note="MLTIC repeat: matches 420. .463 of consensus"
repeat_region	17036..17338
repeat_region	/note="AluJo repeat: matches 1. .301 of consensus"
repeat_region	17339..17730

```
/note="MUT1C repeat: matches 1. .420 of consensus"
17958..18269
/note="AluX repeat: matches 1. .311 of consensus"
18288..18315
/note="MER91 repeat: matches 32. .59 of consensus"
18383..18507
/note="MIR repeat: matches 23. .148 of consensus"
18544..18650
/note="MIR repeat: matches 137. .261 of consensus"
18799..18820
/note="11 copies 2 mer aa 100% conserved"
18835..19184
/note="THE1A repeat: matches 1. .354 of consensus"
19231..19522
/note="match: STS: Em:HS153YF5"
19297..19346
/note="25 copies 2 mer ca 100% conserved"
19426..19564
/note="MIR repeat: matches 6. .147 of consensus"
19624..19916
/note="AluX repeat: matches 1. .295 of consensus"
20005..20204
/note="MER20 repeat: matches 1. .218 of consensus"
20215..20317
/note="MER5A repeat: matches 4. .108 of consensus"
20410..20541
/note="MER5B repeat: matches 1. .120 of consensus"
20544..20687
/note="AluSg/x repeat: matches 154. .297 of consensus"
20689..20814
/note="FLAM A repeat: matches 14. .136 of consensus"
20816..21218
/note="Charlie4a repeat: matches 17. .427 of consensus"
21337..21634
/note="AluY repeat: matches 1. .297 of consensus"

Query Match      92.1%; Score 1387.2; DB 9; Length 128871;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1470; Conservative 3; Mismatches 32; Indels 6; Gaps 6;

QY 2 CAGGTTCGGGATGACTCTTCTGCTCAATTCGAACACTCATTCATGAGCAAGAGC 61
DB |||||
QY 62 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGCATTTGGAACTTATGCTTGT 121
DB |||||
QY 71523 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGCATTTGGAACTTATGCTTGT 1464
DB |||||
QY 122 TTGTGAAAGAACCAAGTGAACATCACTGAGCTTCTTAAAGTTGCAAGAAAGTTAGAGG 181
DB |||||
QY 71463 TTGTGAAAGAACCAAGTGAACATCACTGAGCTTCTTAAAGTTGCAAGAAAGTTAGAGG 71404
DB |||||
QY 182 ACTATACACTTTCTTTTGAACCTTTATATAATATTTGCTCTGG-TTTTGAAACCCAGG 240
DB |||||
QY 71403 ACTATACACTTTCTTTTGAACCTTTATATAATATTTGCTCTGGTTTGGAAACCCAGG 71344
DB |||||
QY 241 ACTGTTTGA-GGTGAGTGACAGGCTTAC-AGTGGCCTTTAATCCAATCCAGAAATTCG 298
DB |||||
QY 71343 GCTGTTAGAGGGTGAGTGACAGCTTACAGTGGCCTTATTCNACTCCAGAAATTCG 71284
DB |||||
QY 299 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 358
DB |||||
QY 71283 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 71224
DB |||||
QY 359 CTCCTTAATGTACATGAT-GGCCAAGAGTATGGCAGCTCTCTTGGCAGTCCGATGGAG 417
DB |||||
QY 71223 CTCCTTAATGTACATGATGGCCAGAAAGTATGGCAGCTCTCTTGGCAGTCCGATGGAG 71164
DB |||||
QY 418 ATGGA-GATGCTTGTCAATGAAGGGCCNCTGTTGTCAATTCGAGCTACACAAGAA 476
DB |||||
QY 71163 ATGGAGAGTGCCTTGTCAATGAAGGGCCGCTGTTGTTCATTCGAGCTACACAAGAA 71104
DB |||||
QY 477 AAAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCTTGCGCCAGAC 536
```

RESULT 7

```
DB 71103 AAAAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCGCCAGAC 71045
QY 537 CTTTACACATTCAGAAAGACCTTAATAAACAATGCTTAATGCAACACCGGCTTACCTCTG 596
DB 71044 CTTTACACATTCAGAAAGACCTTAATAAACAATGCTTAATGCAACACCGGCTTACCTCTG 70985
QY 597 TGAACAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGT 656
DB 70984 TGAACAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGT 70925
QY 657 GCGACAGAACCTCCCAAGGAAAGAAATGCAAGGAAATGAATTTAGCTGTGAGGTATG 716
DB 70924 GCGAACAGAACCTCCCAAGGAAAGAAATGCAAGGAAATGAATTTAGCTGTGAGGTATG 70865
QY 717 TGGGCGAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAGAAACACACAAGATTTC 776
DB 70864 TGGGCGAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAGAAACACACAAGATTTC 70805
QY 777 TTTCACTTACGGGTTAAACATGTCGGAAAGAAATTCAGAGAGCTTGTCTTTTAAAAA 836
DB 70804 TTTCACTTACGGGTTAAACATGTCGGAAAGAAATTCAGAGAGCTTGTCTTTTAAAAA 70745
QY 837 TCACATGCGGACRCAATATGCGAAATCGGGGCCAGAGCAAACTGCAGCAAGGCTTGA 896
DB 70744 TCACATGCGGACRCAATATGCGAAATCGGGGCCAGAGCAAACTGCAGCAAGGCTTGA 70685
QY 897 GAGTAGTCAGCAACGATCAACGAGGTGTCAGAGTGTCACGCGCGCGAGAGCATCTCCTC 956
DB 70684 GAGTAGTCAGCAACGATCAACGAGGTGTCAGAGTGTCACGCGCGCGAGAGCATCTCCTC 70625
QY 957 TCCTTGCAAAATCTGCATGGTTTGTGGCTTCTCTATTTCCAAATAAAGAAAGTCTAATTGA 1016
DB 70624 TCCTTGCAAAATCTGCATGGTTTGTGGCTTCTCTATTTCCAAATAAAGAAAGTCTAATTGA 70565
QY 1017 GCACGCAAGGTGTCACACCAAAAAAATCTGCTTTTCGTACACAGCGCGCGAGAGCATCTC 1076
DB 70564 GCACGCAAGGTGTCACACCAAAAAAATCTGCTTTTCGTACACAGCGCGCGAGAGCATCTC 70505
QY 1077 TCACAAGGAGAAATGCGGTCTCGAGGAGGACTTCTCGCAGTGTGTTCAACTTGAGACC 1136
DB 70504 TCACAAGGAGAAATGCGGTCTCGAGGAGGACTTCTCGCAGTGTGTTCAACTTGAGACC 70445
QY 1137 AAAATCTCACCTCGAAACGGGGAAGAGCTGTCCAGATGTCATCCCTCAGCTCGATCCGTT 1196
DB 70444 AAAATCTCACCTCGAAACGGGGAAGAGCTGTCCAGATGTCATCCCTCAGCTCGATCCGTT 70385
QY 1197 CACCACTTTCAGGCTTGGCAKCTGCTTACAAAGAAWAGTTGCCATTTGCAAGAAAGT 1256
DB 70384 CACCACTTTCAGGCTTGGCAKCTGCTTACAAAGAAWAGTTGCCATTTGCAAGAAAGT 70325
QY 1257 GAAGCAATTTGGGCAAGAGGAGGACACGACAAACGATTCGAGTTCCGAGAGGAGCT 1316
DB 70324 GAAGCAATTTGGGCAAGAGGAGGAGGACACGACAAACGATTCGAGTTCCGAGAGGAGCT 70265
QY 1317 TGGAGAAACAAATAAGAAACCATTTGTGCGGCTCTCGCAAGAGAAAGAGAGTGCAGAAACA 1376
DB 70264 TGGAGAAACAAATAAGGCGAGTTGTGCGGCTCTCGCAAGAGAAAGAGAGTGCAGAAACA 70205
QY 1377 CTTCCACGCGGAGGCGGCTTCCGTGGACGCGATCCCAAGTTACCCAGTATGCAAGAGAA 1436
DB 70204 CTTCCACGCGGAGGCGGCTTCCGTGGACGCGATCCCAAGTTACCCAGTATGCAAGAGAA 70145
QY 1437 GCCCACTCACTGCTCCGAGTGGGCAAGAGCTTTTCAGAACCTACCAAGCTGCTTGA 1496
DB 70144 GCCCACTCACTGCTCCGAGTGGGCAAGAGCTTTTCAGAACCTACCAAGCTGCTTGA 70085
QY 1497 CTTCCAGGCTCC 1507
DB 70084 CTTCCAGGCTCC 70074
```

AR157086	AR157086	5632 bp	DNA	linear	PAT 08-AUG-2001
LOCUS	Sequence 3 from patent US 6242590.				
DEFINITION	Antisense modulation of zinc finger protein-217 expression				
ACCESSION	AR157086				
VERSION	AR157086.1	GI:15125790			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 5632)				
AUTHORS	Cowser, L.M.				
TITLE	Antisense modulation of zinc finger protein-217 expression				
JOURNAL	Patent: US 6242590-A 3 05-JUN-2001;				
FEATURES	Location/Qualifiers				
source	1..5632				
	/organism="unknown"				
	/mol_type="unassigned DNA"				
ORIGIN					
Query Match	88.7%; Score 1337.2; DB 6; Length 5632;				
Best Local Similarity	97.2%; Pred. No. 0;				
Matches 1420; Conservative	3; Mismatches 32; Indels 6; Gaps 6;				
QY	52 GACAAAGAGCACTAATGCTTGTGCTGATTCATATTTGAATCGAGGCAATTTGGAAACCCCTG 111				
DB	1 GACAAAGAGCACTAATGCTTGTGCTGATTCATATTTGAATCGAGGCAATTTGGAAACCCCTG 60				
QY	112 TATGCTTGTGTTGGAAAGAACGATGACACCATCAGCTTCCATAAAGTTTCGAAG 171				
DB	61 TATGCTTGTGTTGGAAAGAACGATGACACCATCAGCTTCCATAAAGTTTCGAAG 120				
QY	172 AAGTTAGAGGACTATACACTTCTTTTGAATTTTATATTAATTAATTTGCTCTGG-TTTT 230				
DB	121 AAGTTAGAGGACTATACACTTCTTTTGAATTTTATATTAATTAATTTGCTCTGGTTTTT 180				
QY	231 GGAAACCCAGGACTGTTAGA-GGGTGAGTGACAGGTCTTAC-AGTGGCCCTTAATCCAACTC 288				
DB	181 GGAAACCCAGGCTGTTAGAGGGTGAGTGACAACTTACAACTGAGTCTTATTCACATC 240				
QY	289 CAGAAATTTGCCCAACGGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 348				
DB	241 CAGAAATTTGCCCAACGGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300				
QY	349 AACTCAATCCCTCTTAATGTACATGAT-GGCCAAGAGTGATTTGCAGCTCTCTTGGCAG 407				
DB	301 AACTCAATCCCTCTTAATGTACATGATGGGCCAAGTGATTTGCAGCTCTCTTGGCAG 360				
QY	408 TCCGATGGAGATGGA-GATGCCCTTGTCAATGAAGGGCCNCCTGTTGTCAATTTCCGAGCT 466				
DB	361 TCCGATGGAGATGGAGGATGCCCTTGTCAATGAAGGGCCCGCTGTTGTTCATTTCCGAGC 420				
QY	467 ACACAAAGAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCT 526				
DB	421 TACACAAAGAAAATGTCTTCGGTGTGAAGCAGAGTATCTCAGTCCGCTTGATA 479				
QY	527 GCAGCCAGACCTTCACACATTGAGAGACCTTAAATAAATGCTCTTAATGCAACACCCGC 586				
DB	480 GCAGCCAGACCTTCACACATTGAGAGACCTTAAATAAATGCTCTTAATGCAACACCCGC 539				
QY	587 CTACCCCTCTGTGAACACGAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATA 646				
DB	540 CTACCCCTCTGTGAACACGAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATA 599				
QY	647 AAAGTCAAGTGGGAACAGAACCTCCCAAGGAAAAGAAATGCAAGAAAATGAATTTAGCT 706				
DB	600 AAAGTCAAGTGGGAACAGAACCTCCCAAGGAAAAGAAATGCAAGAAAATGAATTTAGCT 659				
QY	707 GTGAGGTATGGGCAGACATTTAGAGTGCCTTTTGTGTTGAGATCCACATGAGAACAC 766				
DB	660 GTGAGGTATGGGCAGACATTTAGAGTGCCTTTTGTGTTGAGATCCACATGAGAACAC 719				
QY	767 ACAAGATTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAATTCAGGAGCCCTTGGT 826				

Db	720	ACAAAGATTCTTTCACTTACCGGTGTAA	CATGTGCGGAAGAAGATTCAAGGAGCCTTGGT	779
Qy	827	TTCTTAAAAATCACATGCGGACRCATAAT	TGCGGAGCAAGCAAACTGCAGC	886
Db	780	TTCTTAAAAATCACATGCGGACACATAAT	TGCGGAGCAAACTGCAGC	839
Qy	887	AAGCCTTGGAGAGTAGTCCAGCAACGAT	CAACGAGGTGCTCCAGGTGCAACGCGGCCGAGA	946
Db	840	AAGCCTTGGAGAGTAGTCCAGCAACGAT	CAACGAGGTGCTCCAGGTGCAACGCGGCCGAGA	899
Qy	947	GCATCTCTCTCTTTCGCAAAATTCGATG	TGTTGGCTTCTTATTTCCAAATAAAGAAA	1006
Db	900	GCATCTCTCTCTTTCGCAAAATTCGATG	TGTTGGCTTCTTATTTCCAAATAAAGAAA	959
Qy	1007	GTCTAATTTGAGCACCGCAAGGTGCACA	CAAAAAAACTGCTTTTCGGTACCAAGGAGCGCGC	1066
Db	960	GTCTAATTTGAGCACCGCAAGGTGCACA	CAAAAAAACTGCTTTTCGGTACCAAGGAGCGCGC	1019
Qy	1067	AGACAGACTCTCCAAAGGAGGAATGCCG	TCTCCAGGGAGGACTTCTCGAGTTGTTCA	1126
Db	1020	AGACAGACTCTCCAAAGGAGGAATGCCG	TCTCCAGGGAGGACTTCTCGAGTTGTTCA	1079
Qy	1127	ACTTGAGACCAAAATCTCACCTCGAAAC	CGGGGAAGACCTGTTCAGATGCATCCCTCAGC	1186
Db	1080	ACTTGAGACCAAAATCTCACCTCGAAAC	CGGGGAAGACCTGTTCAGATGCATCCCTCAGC	1139
Qy	1187	TCGATCCGTTTACCACCTTCCAGGCTTGG	CAKCTGGCTACCAAGGAAGTTGCCATTT	1246
Db	1140	TCGATCCGTTTACCACCTTCCAGGCTTGG	CAGCTGGCTACCAAGGAAGTTGCCATTT	1199
Qy	1247	GCCAAGAGTGAAGAAATTGGGGCAAGAGG	GAGCACCGACAAACGACGATTCGAGTTCCG	1306
Db	1200	GCCAAGAGTGAAGAAATTGGGGCAAGAGG	GAGCACCGACAAACGACGATTCGAGTTCCG	1259
Qy	1307	AGAAGGAGCTTGGAGAAAACAAATTAAG	AACCAATTGTGACAGCGCTCTCGCAAGAGAAGA	1366
Db	1260	AGAAGGAGCTTGGAGAAAACAAATTAAG	AGGCAAGTTGTGACAGCGCTCTCGCAAGAGAAGA	1319
Qy	1367	AGTGCMAACACTCCACGGCGAAGCGCC	TCCGTGACGCGGATCCCAAGTTACCCAGTA	1426
Db	1320	AGTGCMAACACTCCACGGCGAAGCGCC	TCCGTGACGCGGATCCCAAGTTACCCAGTA	1379
Qy	1427	GCAAGGAGAAAGCCACTCACTGCTCCG	AGTGGCGCAAGCTTTTCAGAACCTACCAACG	1486
Db	1380	GCAAGGAGAAAGCCACTCACTGCTCCG	AGTGGCGCAAGCTTTTCAGAACCTACCAACG	1439
Qy	1487	TGGTCTTGCACTCCAGGGTCC	1507	
Db	1440	TGGTCTTGCACTCCAGGGTCC	1460	
RESULT 8				
AF041259				
LOCUS				
DEFINITION				
Homo sapiens breast cancer putative transcription factor (ZABC1)				
mRNA, complete cds.				
ACCESSION				
AF041259				
VERSION				
AF041259.1				
KEYWORDS				
GI:3335396				
SOURCE				
Homo sapiens (human)				
ORGANISM				
Homo sapiens				
REFERENCE				
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
1 (bases 1 to 5632)				
AUTHORS				
Collins, C., Rommens, J.M., Kowbel, D., Godfrey, T., Tanner, M.,				
Hwang, S.-I., Polikoff, D., Nonet, G., Cochran, J., Myambo, K.,				
Jay, K.E., Froula, J., Cloutier, T., Kuo, W.-L., Yaswen, P., Dairkee, S.,				
Giovannola, J., Hutchinson, G.B., Isola, J., Kallioniemi, O.-P.,				
Palazzo, M., Martin, C., Ericsson, C., Pinkel, D., Albertson, D.,				
Li, W.-B. and Gray, J.W.				
TITLE				
Positional cloning of ZNF217 and NAB1: genes amplified at 20q13.2				
and overexpressed in breast carcinoma				
JOURNAL				
Proc. Natl. Acad. Sci. U.S.A. 95 (15), 8703-8708 (1998)				

MEDLINE 98337979
PUBMED 9671742
REFERENCE 2 (bases 1 to 5632)
AUTHORS Hwang, S.-I., Polikoff, D., Nonet, G., Cochran, J., Myambo, K., Jay, K.E., Proula, J., Cloutier, T., Kuo, W.-L., Yaswen, P., Dairkee, S., Giovanola, J., Hutchinson, G.B., Isola, J., Kallioniemi, O.-P., Palazzolo, M., Martin, C., Ericsson, C., Pinkel, D., Albertson, D., Li, W.-B. and Gray, J.W.
TITLE Direct Submission
JOURNAL Submitted (07-JAN-1998) Life Sciences, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA
FEATURES
source Location/Qualifiers
1..5632
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20q13.2-q13.3"
1..5632
/gene="ZABC1"
272..3418
/gene="ZABC1"
/note="Kruppel-like"
/codon_start=1
/product="breast cancer putative transcription factor"
/protein_id="AAC39895.1"
/db_xref="GI:3335397"
/translation="MQSKVTGNMPTQSLIMYMDGPEVIGSSISGSPWEMEDALSMDGTA
VVPFRATQENKVIQIEGYMPLDQMFCSQTFTHSEDLNKHVLMQHRPTLCEPAVLVEA
EYLPDQKQVREPPKRNKTHNGKSGARSKLQGLSESPATINEVQVHAESISSPYKIC
CGRFPKBPFLKTHNGKSGARSKLQGLSESPATINEVQVHAESISSPYKIC
MYCGFLFPNKESLIEHRKVTHTKTAFGTSSAQTDQPGMPSSREDFLQFLNRPKSH
PTGKKPVCIPOLDPFTTFQAWLATKGVKAIQVQKESGQESDNDSDSSKEELG
ETNKGSCAGLSEKESKSHGEAPSDADPKLPSSKEKPTHCSECGAFRTYHOLVL
HSRVKDRAGAESPTKSHVSDGROPTCSPDLAAPLDENGAVDRGEGSEDSERGLP
EGHLDNDGGRKIRLHLSRESCYCGKFRSNYLNLHRTHTGKPKYKCFBCTYAA
AOKTSLRHLRHHKEKQTDVAAEVKNDQKQDTEALLTADSAQTKNLKRPFDGAKD
VTGSPKQLKEMPSVFNVLGSAVLSPAHKQTDQPKNAADSDADKYNKNPFPAYLD
LLKRSASVETQANNLIRTKADVTPPDGCTTHNLEVPKEQKQETETAADCRVPSVDC
HKPLNLSVGLHNCPAISLSKSLIPSIITCPCTKTFYFVPLVMHQRLEHKNVNDVH
KNCRNLSLRRTGCPALLGKDVPLPSFCFKPKPSAPPAQSKLSPSAKGQSPFG
PKATKLPVGSSTLAPSLNKHSHRPOQNVGGAATRQOQSEMFPTKSVSPADPKTKR
PETKLPKLVAPQPTLGSNSNGSIDYPAKNSWPAPPRDYFCNRSASNTAAEFGE
PLPKLKSQVVALVDQPCANYRRGYDLPKYHVRGITSLLPQDCVYPSQALPPKPRF
LSSSEVDSPNVLTVOKPYGSGPLYTCVPAGSPASSSTLEGKRPVSYQHLNSMAQKR
NYENFIGNAHYRPNDKKT"
470..535
/gene="ZABC1"
/note="zinc finger 1"
659..721
/gene="ZABC1"
/note="zinc finger 2"
744..805
/gene="ZABC1"
/note="zinc finger 3"
923..985
/gene="ZABC1"
/note="zinc finger 4"
1400..1463
/gene="ZABC1"
/note="zinc finger 5"
1688..1750
/gene="ZABC1"
/note="zinc finger 6"
1772..1834
/gene="ZABC1"
/note="zinc finger 7"
2405..2470
/gene="ZABC1"
/note="zinc finger 8"

Query Match 88.7%; Score 1337.2; DB 9; Length 5632;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1420; Conservative 3; Mismatches 32; Indels 6; Gaps 6;
QY 52 GACAAAGAGCAGCTAATGCTTTGCTGCTGATTCATATTTGAATCGAGGCAATTTGGAAACCCCTG 111
DB 1 GACAAAGAGCACTAATGCTTTGCTGCTGATTCATATTTGAATCGAGGCAATTTGGAAACCCCTG 60
QY 112 TATGCTTCTTTTGTGGAAAGAACCAAGTACACCATCAGGCTTCTTAAAGTTTCAAG 171
DB 61 TATGCTTCTTTTGTGGAAAGAACCAAGTACACCATCAGGCTTCTTAAAGTTTCAAG 120
QY 172 AAGTTAGAGGACTATACACTTTCTTTGAACTTTTATATAATAATATTTGCTCTGG-TTTT 230
DB 121 AAGTTAGAGGACTATACACTTTCTTTGAACTTTTATATAATAATATTTGCTCTGGTTTTT 180
QY 231 GGAACCCAGGACTGTTTGAAG-GGCTGAGTGACAGGCTCTTAC-AGTGGCCCTTAATCCAATC 288
DB 181 GGAACCCAGGCTGTTTGAAGGGGTGAGTGACAGGCTCTTACAGGTCCTTATTTCCAATC 240
QY 289 CAGAAATTGCCCAACCGAACTTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 348
DB 241 CAGAAATTGCCCAACCGAACTTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300
QY 349 AACTCAATCCCTTAAATGTATCATGAT-GGCCAAGAGTGATTTGGCAGCTCTCTTCCAG 407
DB 301 AACTCAATCCCTTAAATGTATCATGATGGCGGAGAAATGATTTGGCAGCTCTCTTGGCAG 360
QY 408 TCCGATGGAGATGGA-GATGCCCTTGTCAATGAAGGGCCNCCTGTTGTCAATTTCCGAGCT 466
DB 361 TCCGATGGAGATGGAGATGCGCTTGTCAATGAAGGGAGCGCTGTTGTTCATTTCCAGC 420
QY 467 ACACAAAGAAAAAATGTCAATCCGAATCGAGGGGAATAATGCGCTTTGGATTGCAATGTTCT 526
DB 421 TACACAGAAAAAATGTC-ATCCAAATCGAGGGGTATATATGCGCTTTGGATTGCAATGTTCT 479
QY 527 GCAGCAGACCTTACACATTCAGAAAGCTTAAATAACATGCTTTAAATGCAACACCGGC 586
DB 480 GCAGCAGACCTTACACATTCAGAAAGCTTAAATAACATGCTTTAAATGCAACACCGGC 539
QY 587 CTACCTCTGTGAACCCAGCAGTCTTTCGGGTGGAAGCAGAGTATCTCAGTCGCTTGATA 646
DB 540 CTACCTCTGTGAACCCAGCAGTCTTTCGGGTGGAAGCAGAGTATCTCAGTCGCTTGATA 599
QY 647 AAAGTCAAGTGCAGAACAGAACCTCCCAAGAAAAAGAAATTCGAAGAAAAATGAATTTAGCT 706
DB 600 AAAGTCAAGTGCAGAACAGAACCTCCCAAGAAAAAGAAATTCGAAGAAAAATGAATTTAGCT 659
QY 707 GTGAGGTATGTGGGAGACATTTAGAGTCGCTTTTGTATGTTGAGATCCACATGAGAACAC 766
DB 660 GTGAGGTATGTGGGAGACATTTAGAGTCGCTTTTGTATGTTGAGATCCACATGAGAACAC 719
QY 767 ACAGAGATTTCTTTCACTTACGGGTGTAAACATGTGCGGAAGAAAGATTCAGAGGAGCTTGT 826
DB 720 ACAGAGATTTCTTTCACTTACGGGTGTAAACATGTGCGGAAGAAAGATTCAGAGGAGCTTGT 779
QY 827 TTCTTTAAAAATCACATGCGGACRCATATGCGCAAAATTCGGGGGCCAGAGCAAAATGTCAGC 886
DB 780 TTCTTTAAAAATCACATGCGGACRCATATGCGCAAAATTCGGGGGCCAGAGCAAAATGTCAGC 839
QY 887 AAGGCTTGAGAGTAGTTCAGCAACATCAACAGAGTCTCCAGGTGACGGCGCCGAGA 946
DB 840 AAGGCTTGAGAGTAGTTCAGCAACATCAACAGAGTCTCCAGGTGACGGCGCCGAGA 899
QY 947 GCATCTCTCTCTTTCGCAAAATCTCATGTTTGTGGCTTCTTATTTCCAAATAAAGAAA 1006
DB 900 GCATCTCTCTCTTTCGCAAAATCTCATGTTTGTGGCTTCTTATTTCCAAATAAAGAAA 959
QY 1007 GTCTAATTTAGCAGCGCAGAGGTGCACACCAAAAAAATCTCTTTTCGGTACAGCGCGC 1066
DB 960 GTCTAATTTAGCAGCGCAGAGGTGCACACCAAAAAAATCTCTTTTCGGTACAGCGCGC 1019
QY 1067 AGACAGACTCTCCCAAGAGGAGAAATCGCGCTCTCGAGGGAGGAGCTTCTTCGCAATTTCA 1126

ORIGIN

Db	1020		AGACAGACTCTCCACAGGGAATGCCGTCTCGAGGGAGGACTTCTCGAGTTGTTCA	1079
Qy	1127		ACTTGAGACCAAAATCTCACCTGAAACGGGGAAGAAGCCTGTGCAGATGCATCCCTCAGC	1186
Db	1080		ACTTGAGACCAAAATCTCACCTGAAACGGGGAAGAAGCCTGTGCAGATGCATCCCTCAGC	1139
Qy	1187		TCGATCCGTTTACCACCTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGTTGCCATTT	1246
Db	1140		TCGATCCGTTTACCACCTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGTTGCCATTT	1199
Qy	1247		GCCAAAGAGTGAAGAAATTTGGGGCAAGAGGGAGCACCGAACGACGATTCGAGTTCCG	1306
Db	1200		GCCAAAGAGTGAAGAAATTCGGGGCAAGAGGGAGCACCGAACGACGATTCGAGTTCCG	1259
Qy	1307		AGAAGGAGCTTGGAGAAACAAATAAGAACCAATTTGTGCAGGCTCTCGCAAGAGAAAGAGA	1366
Db	1260		AGAAGGAGCTTGGAGAAACAAATAAGAGGCAAGTTGTGCAGGCTCTCGCAAGAGAAAGAGA	1319
Qy	1367		AGTGCAACACTCCACCGGGAAGCGCCCTCGTGAGCGCGATFCCCAAGTTACCCAGTA	1426
Db	1320		AGTGCAACACTCCACCGGGAAGCGCCCTCGTGAGCGCGATFCCCAAGTTACCCAGTA	1379
Qy	1427		GCAGGAGAGCCCACTCACTGCTCCAGTGGCGCAAGCTTTTCAGAACTTACCAACGAGC	1486
Db	1380		GCAAGGAGAGCCCACTCACTGCTCCAGTGGCGCAAGCTTTTCAGAACTTACCAACGAGC	1439
Qy	1487		TGGTCTTTGCCACTCCAGGGTCC	1507
Db	1440		TGGTCTTTGCCACTCCAGGGTCC	1460
RESULT 9				
LOCUS	CQ726295	CQ726295	Sequence 12229 from Patent WO02068579.	5633 bp DNA linear PAT 03-FEB-2004
DEFINITION	ACCESSION	CQ726295		
VERSION	CQ726295.1	GI:42288937		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
TITLE	Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.			
JOURNAL	Kits, such as nucleic acid arrays, comprising a majority of			
	humanexons or transcripts, for detecting expression and other uses			
FEATURES	thereof			
	Patent: WO 02068579-A 12229 06-SEP-2002;			
source	PE Corporation (NY) (US)			
	Location/Qualifiers			
ORIGIN	1..5633			
	/organism="Homo sapiens"			
Query Match	/mol_type="unassigned DNA"			
	/db_xref="taxon:9606"			
Best Local Similarity	88.7%;	Score 1337.2;	DB 6;	Length 5633;
Matches 1420;	Conservative	3;	Mismatches 32;	Indels 6; Gaps 6;
Qy	52	GACAAAGAGCACTAATGCTTTGCTGCTGATTCATATTTGAATTCGAGGCAATTTGGGAACCCCTG	111	
Db	1	GACAAAGAGCACTAATGCTTTGCTGCTGATTCATATTTGAATTCGAGGCAATTTGGGAACCCCTG	60	
Qy	112	TATGCTTGTGTTGGAAAGAACCAAGTACCACTACCTAGCTTCTTAAAGTTTCGAAG	171	
Db	61	TATGCTTGTGTTGGAAAGAACCAAGTACCACTACCTAGCTTCTTAAAGTTTCGAAG	120	
Qy	172	AAGTTAGAGCACTATACACTTCTTTTGAACTTTTATAATAATTTGCTCTGG-TTTT	230	
Db	121	AAGTTAGAGCACTATACACTTCTTTTGAACTTTTATAATAATAATTTGCTCTGGTTTTT	180	

Qy	231	GGAAACCCAGGACTGTTTAGA-GGGTGAGTGACAGGCTCTTAC-AGTGGCCTTAATCCAACTC	288	
Db	181	GGAAACCCAGGCTGTTAGAGGGTGAGTGACAGTCTTTCAAGTGGCCTTATTTCCAACTC	240	
Qy	289	CAGAAATTTGCCCAACGGAACTTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC	348	
Db	241	CAGAAATTTGCCCAACGGAACTTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC	300	
Qy	349	AATCAATCCCTCTTAATGTACATGGAT-GGCCAAGAGTGAATGGCAGCTCTCTTCCAG	407	
Db	301	AATCAATCCCTCTTAATGTACATGGATGGCCAGAGTGAATGGCAGCTCTCTTCCAG	360	
Qy	408	TCCGATGAGATGGA-GATGCTTGTCAATGAAAGGGCCNCCTGTGTCAATTCGAGCT	466	
Db	361	TCCGATGAGATGAGGATGCTTGTCAATGAAAGGGACCCGCTGTGTTCATTCGAGC	420	
Qy	467	ACAAAGAAAAAATGCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCT	526	
Db	421	TACAAAGAAAAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCT	479	
Qy	527	GCAGCAGACCTTCACACATTCAGAGACCTTAATAACATGTCTTAATGCNAACCCGC	586	
Db	480	GCAGCAGACCTTCACACATTCAGAGACCTTAATAACATGTCTTAATGCNAACCCGC	539	
Qy	587	CTACCTCTGTGAACAGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCGCTTGATA	646	
Db	540	CTACCTCTGTGAACAGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCGCTTGATA	599	
Qy	647	AAAGTCAAAGTCCGAACAGAACCTCCCAAGGAAAAAGATTGCAAGGAAAAATGAATTAAGT	706	
Db	600	AAAGTCAAAGTCCGAACAGAACCTCCCAAGGAAAAAGATTGCAAGGAAAAATGAATTAAGT	659	
Qy	707	GTGAGTATGTTGGCAGACATTTAGAGTCCGCTTTTGATGTTGAGATCCACATGAGAAC	766	
Db	660	GTGAGTATGTTGGCAGACATTTAGAGTCCGCTTTTGATGTTGAGATCCACATGAGAAC	719	
Qy	767	ACAAAGATTCTTTCACTTACCGGTGTAAATGTCGGGAAGAAAGATTCAAGAGCCTTGGT	826	
Db	720	ACAAAGATTCTTTCACTTACCGGTGTAAATGTCGGGAAGAAAGATTCAAGAGCCTTGGT	779	
Qy	827	TTCTTAAATAATCACATGCGGACRCATATATGCAAAATCGGGGCCAGAGCAAACTGCAGC	886	
Db	780	TTCTTAAATAATCACATGCGGACRCATATATGCAAAATCGGGGCCAGAGCAAACTGCAGC	839	
Qy	887	RAAGCTTGGAGAGTAGTCCAGCAACGATCAAGAGTCTCCAGTGCACCGGCCGAGAG	946	
Db	840	RAAGCTTGGAGAGTAGTCCAGCAACGATCAAGAGTCTCCAGTGCACCGGCCGAGAG	899	
Qy	947	GCATCTCTCTCTTGCATAAATCTGCATGTTTGTGGCTTCTCTATTTCCAAATAAGAAA	1006	
Db	900	GCATCTCTCTCTTGCATAAATCTGCATGTTTGTGGCTTCTCTATTTCCAAATAAGAAA	959	
Qy	1007	GTCATAATGAGACCGCAAGGTGCACACCAAAAAAATCTGCTTTTCGGTACAGAGCGGCG	1066	
Db	960	GTCATAATGAGACCGCAAGGTGCACACCAAAAAAATCTGCTTTTCGGTACAGAGCGGCG	1019	
Qy	1067	AGACAGACTCTCCACAGGAGGAATGCGCTCTCGAGGGAGACTTCTCGAGATTGTTCA	1126	
Db	1020	AGACAGACTCTCCACAGGAGGAATGCGCTCTCGAGGGAGACTTCTCGAGATTGTTCA	1079	
Qy	1127	ACTTGAGACCAAAATCTCACCTCGAAACGGGGAAGAACCTGTTCAGATGCAATCCCTCAGC	1186	
Db	1080	ACTTGAGACCAAAATCTCACCTCGAAACGGGGAAGAACCTGTTCAGATGCAATCCCTCAGC	1139	
Qy	1187	TCGATCCGTTCCACCTTCCAGGCTTGGCAKCTGGCTTACCAAGGAAGTTGCCATTT	1246	
Db	1140	TCGATCCGTTCCACCTTCCAGGCTTGGCAKCTGGCTTACCAAGGAAGTTGCCATTT	1199	
Qy	1247	GCCAAAGAGTGAAGAAATTTGGGGAAGAGGAGCACCGACAAACGACGATTCGAGTTCCG	1306	
Db	1200	GCCAAAGAGTGAAGAAATTTGGGGAAGAGGAGCACCGACAAACGACGATTCGAGTTCCG	1259	
Qy	1307	AGAAGGAGCTTTGGAGAAAACAAATTAAGAACCATTTGTGAGGCTCTCTCGCAAGAGAAAGAGA	1366	

```
|||||
1260 AGAAGAGCTTGGAGAAACAAATAAGGCGAGTTGTGAGGCTCTCGCAGGAGAAAGAGA 1319
QY
1367 AGTGCAAAACATCCCAACGCGGAGCCCTCGGTGACCGGATCCCAAGTTACCCAGTA 1426
Db
1320 AGTGCAAAACATCCCAACGCGGAGCCCTCGGTGACCGGATCCCAAGTTACCCAGTA 1379
QY
1427 GCAAGGAGAGCCCACTCACTGCTCGAGTGGGCAAAAGCTTTCAGAACTTACCAACAGC 1486
Db
1380 GCAAGGAGAGCCCACTCACTGCTCGAGTGGGCAAAAGCTTTCAGAACTTACCAACAGC 1439
QY
1487 TGGTCTTGCACCTCCAGGTC 1507
Db
1440 TGGTCTTGCACCTCCAGGTC 1460

RESULT 10
BD085734
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITILE
JOURNAL
COMMENT
OS Artificial Sequence
PN JP 2001524802-A/10
PD 04-DEC-2001
PF 15-JUL-1997 JP 1998506264
PR 15-JUL-1996 US 08/680395,16-OCT-1996 US 08/731499 PR
17-JAN-1997 US 08/785532
PI JOE W GRAY, COLIN CONRAD COLLINS, SOO IN HWANG, TONY GODFREY, PI
DAVID KOWBEL,
PI JOHANNA ROMMENS
PC C12N15/11,C12Q1/68,A61K48/00
CC Description of Artificial Sequences:ZABCl Open Reading Frame FH
Key Location/Qualifiers
FT source 1..3186
FT /organism='Artificial Sequence'.
FEATURES
source
1..3186
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'

ORIGIN
Query Match 73.6%; Score 1108.6; DB 6; Length 3186;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 1156; Conservative 3; Mismatches 28; Indels 3; Gaps 3;
QY 320 ATGCAATCGAAGTGCACGAGCAACATGCAACTCAATCCCTCTTAATGATACATGATGG 378
Db 1 ATGCAATCGAAGTGCACGAGCAACATGCAACTCAATCCCTCTTAATGATACATGATGG 60
QY 379 CCAAGAGTATGGGAGCTCTCTTGGCAGTCCGATGGAGATGGA-GATGCCCTTGTCAATG 437
Db 61 CCAGAAGTATGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAATGCTTGTCAATG 120
QY 438 AAAGGCCNCCTGTTGTCAATTCGAGCTACACAAGAAAAAATGTCAATCCGATCGA 497
Db 121 AAAGGACCGCTGTTGTTCTTCATTCGAGCTACACAAGAAAAAATGTCT-ATCCAATCGA 179
QY 498 GGGGATATGCCCTTGGATGATGTTCTGAGCCAGACCTTCACACATTCAGAGACCT 557
Db 180 GGGGATATGCCCTTGGATGATGTTCTGAGCCAGACCTTCACACATTCAGAGACCT 239
```

```
QY 558 TAATAAACAATGCTCTTAATGCAACACACGCGCTACCTCTGTGTGAAACAGCAGTCTTCCGGT 617
Db 240 TAATAAACAATGCTCTTAATGCAACACACGCGCTACCTCTGTGTGAAACAGCAGTCTTCCGGT 299
QY 618 TGAACGACAGTATCTCAGTCCGCTTGTATAAAGTCAAGTGCAGGCAACAGAACCTCCCAAGA 677
Db 300 TGAACGACAGTATCTCAGTCCGCTTGTATAAAGTCAAGTGCAGGCAACAGAACCTCCCAAGA 359
QY 678 AAAGAATTCGAAGGAAATGAATTTAGTGTGAGGTATGTGGGCGAGACATTTAGAGTCGC 737
Db 360 AAAGAATTCGAAGGAAATGAATTTAGTGTGAGGTATGTGGGCGAGACATTTAGAGTCGC 419
QY 738 TTTTGAATGTTGAGATCCACATGAGAACACACAAAGATTTCTTCACTTACGGGTGTAACAT 797
Db 420 TTTTGAATGTTGAGATCCACATGAGAACACACAAAGATTTCTTCACTTACGGGTGTAACAT 479
QY 798 GTGCGGAGAAAGATTCAGGAGCCTTGTCTTAAATAATCACATGCGGACATCAATG 857
Db 480 GTGCGGAGAAAGATTCAGGAGCCTTGTCTTAAATAATCACATGCGGACATCAATG 539
QY 858 CAAATCGGGGCGCAGAGCAAACTGACGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAA 917
Db 540 CAAATCGGGGCGCAGAGCAAACTGACGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAA 599
QY 918 CGAGTGTCTCAGGTGCGCGGCGGAGAGCATCTCTCTTCTTGCATAAATCTGCATGTT 977
Db 600 CGAGTGTCTCAGGTGCGCGGCGGAGAGCATCTCTCTCTTCAAAAATCTGCATGTT 659
QY 978 TTGTGCTCTCTTATTTCCAAATAAAGAAAGTCTTAATTGAGCAGCCGCAAGTGCACACCA 1037
Db 660 TTGTGCTCTCTTATTTCCAAATAAAGAAAGTCTTAATTGAGCAGCCGCAAGTGCACACCA 719
QY 1038 AAAAACTGCTTTCGGTACACGAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTC 1097
Db 720 AAAAACTGCTTTCGGTACACGAGCGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTC 779
QY 1098 CTCGAGGAGAGTCTCTGAGTGTTCACATTGAGACCAAAATCTCACCCTGAAACGGG 1157
Db 780 CTCGAGGAGAGTCTCTGAGTGTTCACATTGAGACCAAAATCTCACCCTGAAACGGG 839
QY 1158 GAAGAGCTGTGATGATCCCTCAGTCGATCCGTTCCCAAGAGTGAAGGAATTCGGGCAAGAG 1217
Db 840 GAAGAGCTGTGATGATCCCTCAGTCGATCCGTTCCCAAGAGTGAAGGAATTCGGGCAAGAG 899
QY 1218 KCTGGCTACCAAGGAAGTTCGCAATTTGCCAAGAGTGAAGGAATTCGGGCAAGAG 1277
Db 900 GCTGGCTACCAAGGAAGTTCGCAATTTGCCAAGAGTGAAGGAATTCGGGCAAGAG 959
QY 1278 GAGCACCGACACGACGATTCGAGTTCGAGAGGAGCTTTGGAGAAACAAATAAGAACCA 1337
Db 960 GAGCACCGACACGACGATTCGAGTTCGAGAGGAGCTTTGGAGAAACAAATAAGGGCAG 1019
QY 1338 TTGTGAGGCTCTCGAAGAGAGAGAGTGAACACTCCACGCGGAGAGCCCTC 1397
Db 1020 TTGTGAGGCTCTCGAAGAGAGAGAGTGAACACTCCACGCGGAGAGCCCTC 1079
QY 1398 CQTGACGCGGATCCCAAGTACCCAGTAGCAAGGAGAGCCCACTCACTGCTCCGAGTG 1457
Db 1080 CQTGACGCGGATCCCAAGTACCCAGTAGCAAGGAGAGCCCACTCACTGCTCCGAGTG 1139
QY 1458 CGGCAAGCTTTCAGAACCTTACCAACAGCTGGTCTTGCATCTCCAGGGTCC 1507
Db 1140 CGGCAAGCTTTCAGAACCTTACCAACAGCTGGTCTTGCATCTCCAGGGTCC 1189

RESULT 11
AC118373
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
AC118373
Rattus norvegicus clone CH230-35619, *** SEQUENCING IN PROGRESS
***, 6 unordered pieces.
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
```


Db 139495 GCTTCTGGAGGTTTGAGTGGTGATATCTCACAAGTGTCTTATTTAACTCCAGAGATCC 139554
QY 298 CCCAAGCGAATCTTGAGATTATATGCA-----ATCGAAAGTGCAGAGAAACATGCCAACT 352
Db 139555 TTAGCAGACAGCTGGATCGTCCATCACTCAGACCTGAGTGGCGGAGCATGCGACG 139614
QY 353 CAATCCCTCTTAATGTACATGGA-TGGCCAAAGATGATTGGCGAGCTCTCTTGTCCAGTCCG 411
Db 139615 CAATCCCTCTCGTGTACATGGAACGCGCGGAGTCTCTTGGCAGTCTCTTAGTATCCCA 139674
QY 412 ATGGAGATGGA-GATGCGCTTGTCAATGAAGAGGCCNCTGTTGTCAATTCGAGCTTAC 470
Db 139675 ATGGAGGTGATGATGCTGTGCCCATGAAGGACCGGCTGTGTCCTCCCTTC-CGAGCTGC 139733
QY 471 AAAGAAAAAATGTCAATCCGAATCAGAGGGGAATATGCCCTTTGGATTGTCATGTTCTGCAG 530
Db 139734 GCAGGAGAGAGCTTGGCCACAGCTCAGGAGCACATGCGCCCTGGACTGCATGTTCTGCAG 139793
QY 531 CCAGACCTTTCACATTCAGAGACCTTAAATAAACATGTCTTAATGAACACACCGGCTTAC 590
Db 139794 CCAGGTCTTCTCTCAGCGGAAGATCTCAATCAGCATGTGCTCTCTGCAGCAGCAGCGCCAC 139853
QY 591 CTTCTGTGAACAGCAGTCTTCTCGGCTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAG 650
Db 139854 CTTCTGTGAGCGGCTGTGCTGCTGTGAGGCGCGAGTATTTAAGCCCTCGACAAAGG 139913
QY 651 TCAAGTCCGAACAGAACCTCCCAAGGAAAGAAATTCGAAGGAAA-----ATGAATTTAGCTG 707
Db 139914 TCAGGTGCCACCGAACCAACCAAGGAGAGAGCAGCAAGAAATGGAAGAGCTGAGCTG 139973
QY 708 TGAGGTATGTGGGAGACATTTAGAGTCGCTTTTGTGATGTGTGATGCCATGAGAACACA 767
Db 139974 TGATGTATCGGACACACATTCCTCAGTGGCTTTTGTGATGTGTGAGGCCACATGAAGAAA 140033
QY 768 CAAAGATCTTCTCACTTAGCGGTGTAAATGTCGCGAAGAAATTCAGAGGACCTTGGTT 827
Db 140034 TAAGGACTCTTTCAGCTATGTGGGTGACGATGTGCGGAGGAGATTCAGAGGACCGTGGTT 140093
QY 828 TCTTAAAAATCATATCGGACRCATATATGGCAAAATCGGGGGCAGAGCAAACTGCAGCA 887
Db 140094 CTGGAAGACCATGCGGACACACATATGCGAAGTCTGSCACCAAGAGCAAGCTGCAGCA 140153
QY 888 AGGCTTGGAGATGTTCAGCAACGATCAACAGAGTGTCTCAGGTGCACGCGGCGAGAG 947
Db 140154 GGGCAGGAG---AGTCCAGTCAACATCAACGATGTGTCTCAGGACATGCTCCCTGGGAA 140210
QY 948 CATCTCTCTCTTGCAGAAATCTGCATGTTGTGGCTTCTTATTTCCAAATGAAGAAG 1007
Db 140211 CATCTCCACTCCCTACAAGATCTGCATGTGTGTGGCTTCTCTTCCCAATGAAGCAGAG 140270
QY 1008 TCTAATTGAGCACCGCAAGGTGCACACCAAAAAAACTGCTTTCCGTTACCAGCAGCGCGCA 1067
Db 140271 CTCTATTGAGCACAGAGGTTTCATGCCAAGAAATGTTCTTAGTGCCAGCAGCAGCAGC 140330
QY 1068 GACAGATCTTCCAAAGGAGGATCCGCTCTCAGGAGGAGACTTCTCTGAGTGTGTCAA 1127
Db 140331 CCCTGATGCCCAAGAGAGAACCCACATCCCGAGGAGAGAGTCTCTCAGTGTGTTTGA 140390
QY 1128 CTTGAGACCAAAATCTCACCTTGAACGGGAGAGAGCTGTCTCAGATGATCCCTCAGCT 1187
Db 140391 CTTGAGACCAAGATCAACTCTGACAGAGATGACGTGAAGGCCCATGACCTCTATACCTCAGCT 140450
QY 1188 CGATCCGTTTCCACCTCTCAGGCTTGGCAKCTGGCTTACCAAGGAAGTGTGCCATTG 1247
Db 140451 TGACCAATTCACCACTTACAGGATGAGGAGTGGTGTGCTACCAAGGCAAGTGGCGTTGC 140510
QY 1248 CCAAGAA---GTGAAGGAATTTGGGCAAGAGGAGCACCGACCAACGACGATTCGAGTTTC 1304
Db 140511 CCAAGAAAGAGGTGAAGAGTTCAGGCCCAAGAGGAGCACCGACCAACGATGATTCGTTTC 140570
QY 1305 CGAGA---AGGAGCTTGGAGAA----- 1323
Db 140571 AGAGAAAGAGGAACTTTGGAGAAATATGTTGGTGGGGTGAAGGTGGAAGGTTCTCGAAAGTC 140630

QY 1324 ----ACAAATAAGAACCATTTGTGCGGCTCTCGCAAGAGAAAGAGAGTGCAAACACTC 1379
Db 140631 CAAGACAAATAGAACGAGTGTTCAGTCTCTCCCAAGACAAAGATAAGCTAGACATGG 140690
QY 1380 CCACGCGAGAGCGCTCTCGTGGAGCGGATCCCAAGTTTACCAGTAGCAGGAGAGCC 1439
Db 140691 TAAACAGTGAAGTGCCTTTCAGGGGACGCTGACTCCAAAGTTGCCCGGAGCAAGAGAGCC 140750
QY 1440 CACTCACTCTCCGAGTGGCGAAAGCTTTTCAGAACCTACCAACAGCTGTCTTGTGACTC 1499
Db 140751 CACACACTGTTCCGAGTGCATTAAGCGTTTTCAGGACATACCAACAGCTGTCTGACTC 140810
QY 1500 CAGGCTCC 1507
Db 140811 GAGGCTGC 140818

RESULT 12
AC099079
LOCUS
DEFINITION
AC099079 246332 bp DNA linear HTG 10-MAY-2003
Rattus norvegicus clone CH230-99J9, *** SEQUENCING IN PROGRESS ***
6 unordered pieces.
AC099079
HTG; HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.
Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
ORGANISM
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 246332)
Muzny, D. Marie., Metaker, M. Lee., Abramzon, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowls, C., Kraft, C. L., Lebow, H., Levay, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshewa, L., Loulseged, H., Lozado, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Naif, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwakoelameh, O., Okwuonu, G., Olarnpunaagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Fu, L., L.,
Puzo, M., Quirroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartbeyn, A., Sisson, I., Sitter, C. D., Smajls, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,

Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleciyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.

Direct Submission

TITLE
JOURNAL
REFERENCE

Unpublished

2 (bases 1 to 246332)

Worley, K.C.

Direct Submission

TITLE
JOURNAL

Submitted (09-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 246332)

REFERENCE
AUTHORS

Rat Genome Sequencing Consortium.

Direct Submission

TITLE
JOURNAL

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GPZC

Center clone name: CH230-99J9

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 209040 bases at least Q40

Consensus quality: 213312 bases at least Q30

Consensus quality: 216534 bases at least Q20

Estimated insert size: 219845; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 6 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 14481: contig of 14481 bp in length

* 14482 14581: gap of unknown length

* 14582 41011: contig of 26430 bp in length

* 41012 41111: gap of unknown length

* 41112 239819: contig of 198708 bp in length

* 239820 239919: gap of unknown length

* 239920 241122: contig of 1203 bp in length

* 241123 241222: gap of unknown length

* 241223 242316: contig of 1094 bp in length

* 242317 242416: gap of unknown length

* 242417 246332: contig of 3916 bp in length.

* Location/Qualifiers

FEATURES

source 1. 246332
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-99J9"
1. 4672
misc_feature /note="wgs_contig"
136991..139129
misc_feature /note="wgs_contig"
174114..175268
misc_feature /note="wgs_contig"
ORIGIN

Query Match 40.8%; Score 614.2; DB 2; Length 246332;
Best Local Similarity 68.8%; Pred. No. 4.2e-177;
Matches 1079; Conservative 3; Mismatches 397; Indels 89; Gaps 14;

QY 2 CAGGTTGCTGGGATTCAGCTTCTTGCTCAATTGAACACCTCATTTCAATGGAGACAAGAGC 61
DB 50464 CACATTGCGGGATCAACGCTTGTCTCAACTGAGCCACTCACCAGTGGAGATGACAGA 50523
QY 62 ACTAATGCTTTG-TGCTGATTCATATTTGAATCGAGGCAATTGGGAACCTCTGATCCCTTG 120
DB 50524 ACTGTTGCCCTCATGCTGACTCAGCTCTGAATCACAGCCAGGCGAGCTGGATGCTTG 50583
QY 121 TTTGTGGGAAGAACGATGACACCATCACTGAGTTCCTTAAAGTTTCGAAGATTAGAG 180
DB 50584 TTTGTGCAAGACCA---CCTCTATCACTGAGCTTCCCCAAGCTC-----TG 50629
QY 181 GACTATACACTTCTTTTGAACCTTTTATAATAAATATTTGCTCTGTTTGGAAACCCAGG 240
DB 50630 GAATATACACATCCCTCTGAACCTT-----AGAGTTGCTCTAATTTTGACGCCAGG 50680
QY 241 ACTGTTAGAGGTGAGTGACAGGCTTAC---AGTGGCTTAAATCCAACCTCAGAAATTG 297
DB 50681 GCTTCTGGAGGTTTGAGTGGTGATATCTCAAGGTGCTCTATTCTTAATCCAGAGATCC 50740
QY 298 CCCAACGGACTTTGAGATTATATGCA-----ATCGAAGTGACAGGAAACATGCCAACT 352
DB 50741 TTAGCAGACAGCTGGATCGTCCATCCACTGAGCTGAAGTGGCCGCGAGCATGCCAG 50800
QY 353 CAATCCCTCTTAATGTACATGGA-TGGCCAAAGATGATTGGCAGCTCTCTTGCCAGTCCG 411
DB 50801 CAATCCCTCTCGTGTACATGGAACGGCGGGAAGTCTTTGGCAGTTCTTAGTAGCCAG 50860
QY 412 ATGAGATGGA-GATGCTTGTCAATGAAGAGCCCNCTGTTGTCAATTCGAGTACAC 470
DB 50861 ATGAGGTGGATGATGCTGTGCCATGAAGGACCGGCTGTGGTCCCTTC-CGAGCTGC 50919
QY 471 AAAGAAAAAATGTCAATCCGAATCGAGGGAATATGCCCTTGGATTGCATGCTTTCGAG 530
DB 50920 GCAGGAAGAGAGCTTGGCCACAGCTGAGGGAACATGCCCTTGGACTGCTATGTTCTGAG 50979
QY 531 CCAGACCTTTCACACATTCAGAAAGCTTTAATAAATCATGCTTTAATGCAACACCGGCCTAC 590
DB 50980 CCAGTCTTCTCTCAGCGGGAAGATCTCAATCAGCATGTGCTCCTCGAGCACCGGCCAC 51039
QY 591 CCTCTGTGAACACAGAGTTCCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAG 650
DB 51040 CCTCTGTAGCGCGGCTGTGCTGCGTGTGGAGCCGAGTATTTAAGCCCTTCGACAAAG 51099
QY 651 TCAGTTCGGAACAGAACCTCCCAAGGAAAGAAATTTGCAAGGAAA---ATGAATTTAGCTG 707
DB 51100 TCAGTTCGCGGGAACCAACCAAGGAGAGAGCAGCAAGAAATGGAGAGCTGAGCTG 51159
QY 708 TGAGGTATGCGGAGACATTTAGAGTCGCTTTTGTATGTTGAGATCCATGAGAACACA 767
DB 51160 TGATGATGCGGACAGACATTTCCAGTGGCTTTTGTATGTTGAGAGCCACATGAAGAAACA 51219
QY 768 CAAGATTTCTTCACTTACGGGTGTAACTGTGCGGAAGAGATTTCAAGAGCCCTTGGTT 827
DB 51220 TAAGGACTCTTTCACGTATGGGTGCGATGTGCGGAGGAGATTTCAAGAGGCCGCTGGTT 51279

```
QY 828 TCTTAAATAATCATCGGACRCATATATGGCAAAATCGGGGCGCAGAAAGCAAACTGCAGCA 887
Db 51280 CTTGAAGAACCACATCGGACACACATATGCAAGTCTGGCACCAGCAAGCAAGCTGCAGCA 51339
QY 888 AGGCTTTGGAGAGTAGTCCAGCAACGATCAACAGAGTCGTCCAGATGTTGGTCCAGGCACATGCCCTGGAA 947
Db 51340 GGGCAGGAG---AGTCCAGTACCAATCAACGATGTTGGTCCAGGCACATGCCCTGGAA 51396
QY 948 CATCTCTCTCTCTGCAAAATCTGCAATGTTGTGTGGCTTCTTATTTCCAAATAAAGAAAG 1007
Db 51397 CATCTCCACTCCCTACAAGATCTGCATGCTGTGTGGCTTCTTCTTCCCAATAAGCAGAG 51456
QY 1008 TCTAATTGACACCGCAAGGTGCACACCAAAAACCTGCTTTGGTACAGCAGCGCCGCA 1067
Db 51457 CTTCAATTGACACAGCAAGGTTTCATCCAAAGAAACTGTTCTTAGTGCAGCAGCACAGC 51516
QY 1068 GACAGACTCTCCACAGGAGGATGCGCTCTGAGGGAGGACTTCTCGAGTTGTTCAA 1127
Db 51517 CCTGATGGCCACCAAGAGAACCCACATCCCGAGGGAAGAGCTCTCGAGTTTGGAA 51576
QY 1128 CTTGAGACCAAAATCTCACCTGAAACGGGGAAGAGCCTGTGCAGATGCCTCAGCT 1187
Db 51577 CTTGAGACCAAGTCAACTGCAGAGATACAGTGAAGCCCATGCCTCTATACCTCAGCT 51636
QY 1188 CGATCGGTTACACACTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGAGTTGCCATTGG 1247
Db 51637 TGACCCATTCACCACTTACAGGATGCGAGTGGCTTGGCTACCAAGGCAAGAGTGGCGGTTGC 51696
QY 1248 CCAAGAA---GTGAGGAATTTGGGCAAGAGGAGCAGCAGCAGCAGATTCGAGTTC 1304
Db 51697 CCAAGAAAGGTTGAAGAGTTCAGGCCAAGAGGAAGCAGCAGCAGATGATTCGAGTTC 51756
QY 1305 CGAGA---AGGAGCTTGGAGAA----- 1323
Db 51757 AGAGAAGAGGAACTTGGAGAAATATGGGTTGGGGTAAGTGGAGGTTCTGGAAAGTC 51816
QY 1324 ----ACAAATAGAACCATTGTGAGGCTCTCGCAAGAGAAAGAGAAAGTGCACACCTC 1379
Db 51817 CAAGACAAATAGAAGCAGTGTGTTGAGTCTCTCCCAAGCAAAAGATAAGTCTAGACATGG 51876
QY 1380 CCACGCGAAGCGCCCTCGTGGAGCGGATCCCAAGTTACCAGTACGAGGAGAGGCC 1439
Db 51877 TAACTAGTGAAGTGCCTTCAGGGAGCCTGACTCCAAAGTTGGCCCGCAGCAAGAGAGCC 51936
QY 1440 CACTCACTCTCGAGTGGCGCAAGCTTTCAGAACTTACCAACCTGCTGCTTTCGACTTC 1499
Db 51937 CACACACTGTTGCGAGTGCAGTAAGCGTTTCAGGACATACCAACCTGCTGCTGACTTC 51996
QY 1500 CAGGCTCC 1507
Db 51997 GAGGGTGC 52004
```

```
RESULT 13
AL844576/c
LOCUS
DEFINITION Mouse DNA sequence from clone RP23-233M15 on chromosome 2, complete
sequence.
ACCESSION AL844576
VERSION AL844576.9
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 114285)
Direct Submission
TITLE Submitted (22-JAN-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jan 22, 2003 this sequence version replaced gi:27820135.
Sequence from the Mouse Genome Sequencing Consortium whole genome
```

shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/projects/C_elegans/wormpep RP23-233M15 is from the RPCI-23 Mouse BAC Library

constructed by the group of Pieter de Jong.
For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6.

FEATURES

1..114285
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-233M15"
/clone_lib="RPCI-23"

ORIGIN

Query Match 38.1%; Score 573.6; DB 10; Length 114285;
Best Local Similarity 65.6%; Pred. No. 1.3e-164;
Matches 1028; Conservative 3; Mismatches 448; Indels 88; Gaps 10;

```
QY 2 CAGGTGCTGGAGTGGACTTCTTGCTCAATTCAGAACACTCATTCATGAGACAAAGAGC 61
Db 32266 CACATTGGCGGGATCAACGTCTTGCTCACTGAGCCACTCACAGTGGAGATGGACAGAG 32207
QY 62 ACTAATGCTTTGTGCTGATTTCATATTGAAATCGAGGCATTGGGAACCCCTGTATGCTTGT 121
Db 32206 CTGTTGCCCTCGTGTGACGCACATCTGACTACAGCCTGGGCGAGCTTGGATGCTTGT 32147
QY 122 TTGTGGAAGAACACAGTGCACCATCAGCTCTCTTAAAGTTTCGAAGAAGTTAGAGG 181
Db 32146 TTGTGGAAGAGCCA---TCTCCATCACTGAGCTCTCCCAAGCTCC-----GG 32102
QY 182 ACTATACACTTCTTTTGGAACTTTTATAATAATATTGCTCTGTTTGGAAACCCAGGA 241
Db 32101 AATATACTCGTCCCTCTGACCTT-----AGAGGTGCTCTGATTTTGAACCCAGGG 32051
QY 242 CTGTTAGAGGGTGAG---TGACAGGTCTTTACAGTGGCCCTTAATCCAACCTCCAGAAATTGC 298
Db 32050 CTCCTTAGAGTTCTCAGTGGTGATATCTCACAGGCGGCTTATTTCTAAGTGCAGAGATCC 31991
QY 299 CCAACGGAACCTTTGAGATTATATGCAATCGA-----AAGTCACAGGAACATGCCAACT 352
Db 31990 TCAGCAGATAGTGGATTACCCATCTCTGTGCAGACCCTGGAAGTGGTGGCAGCATGCCGCC 31931
QY 353 CAATCCCTCTTAATGTACATGGA-TGGCCACAGATGATGGCAGCTCTCTTCCAGTCCG 411
Db 31930 CAGTCCCTCTCGTGTATATGGAACGCGGGAAGTCTCTCAGCAGCTCTCTAGGCTCCCGAG 31871
```


* 139803 142239: contig of 2437 bp in length
 * 142240 142339: gap of unknown length
 * 142340 149143: contig of 6804 bp in length
 * 149144 149243: gap of unknown length
 * 149244 154777: contig of 5534 bp in length
 * 154778 154877: gap of unknown length
 * 154878 157115: contig of 2838 bp in length
 * 157116 157815: gap of unknown length
 * 157816 161106: contig of 3291 bp in length
 * 161107 161206: gap of unknown length
 * 161207 168000: contig of 6794 bp in length
 * 168001 168100: gap of unknown length
 * 168101 184300: contig of 16200 bp in length
 * 184301 184400: gap of unknown length
 * 184401 208979: contig of 24579 bp in length.

FEATURES

Source

Location/Qualifiers
 1..208979
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="CT7-510C4"
 /clone_lib="CibcJ7 mouse BAC library"

ORIGIN

Query Match 38.1%; Score 573.6; DB 2; Length 208979;
 Best Local Similarity 65.6%; Pred. No. 1.4e-164;
 Matches 1028; Conservative 3; Mismatches 448; Indels 88; Gaps 10;
 QY 2 CAGGTGCTGGGATTCACCTCTTGTCTCAATTCGAACACTCATTCAATGAGACAAAGAGC 61
 DB 94166 CACATGCGGGATCAACGCTCTGTCTCACTGAGCCACTCACAGTGGAGATGGACAGAG 94225
 QY 62 ACTAATGCTTCTGTCTGATTCATATTTGAATCGAGGCATTGGGAACCTGTATGCTTGT 121
 DB 94226 CTGTGCTCTCTGTCTGAGCACAATCTGACTCACAGCCTGGGCGACTTGGATGCTTGT 94285
 QY 122 TTGTGGAAGAACCAAGTGAACATCAGCTGAGCTTCTTAAAGTTGGAAGAGTTAGAGG 181
 DB 94286 TTGTGCAAGAGCCA---TCTCCATCACTGAGCTCTCCCAAGCTC-----GG 94330
 QY 182 ACTATACACTTCTTTTGAACCTTTTATATAATATTTGCTCTGTTTGGAAACCCAGCA 241
 DB 94331 AATATACTCGTCCCTCTGACCTT-----AGAGGTGCTCTGATTTTGCACCCAGGG 94381
 QY 242 CTGTAGAGGGTGAG---TGACAGGTCTTACAGTGGCCCTTAATCCAACTCCAGAAATTCG 298
 DB 94382 CTCCTTAGAGTTCTGAGTGGTATATCTCAAGCGGCCCTTATCTTCACTGAGAGATCC 94441
 QY 299 CCAACGGAACTTTGAGATTATATGCAATCGA-----AAGTGACAGGAACATGCCAAT 352
 DB 94442 TCAGCAGATAGTGGATTACCCATCTGTTCAGACCCCTGAAAGTGGTGGCAGCATGCCACC 94501
 QY 353 CAATCCCTTTAATGATACATGA-TGGCCAGAGTATGGCAGCTCTCTTCCAGTCCG 411
 DB 94502 CAGTCCCTCTCGTGTACATGACGCGGCCGGAAGTCTCTCAGCAGCTCTCTAGGCTCCCGAG 94561
 QY 412 ATGGAGATGGAGATGCTTGTCAATGAAGGCCCCNCTGTGTGCAATTCGAGCTACACA 471
 DB 94562 ATGGAGGTGGATGCTGTGCCCATGAAGGGCCGGTGGCAGTCCCTTCGAGCTGCT 94621
 QY 472 AAGAAAAAATGTCATCCGAATCGAGGGGAATATGCCCTTGGATGTCATGTTTGCAGC 531
 DB 94622 CAGGAGNAGCAGTGGCCGTGGCAGAGGGCCACATGCCCCCTGGAATGTCATGTTCTGCAGC 94681
 QY 532 CAGACCTTCACATTCAGAAACCTTAATAACATGCTTAAATGCAACACCGGCTTACC 591
 DB 94682 CAGGTCTTCTCTCAGCGGAGGATCTCAGTCAGCAGCTGCTCTGCGAGCACCGGCCAC 94741
 QY 592 CTCCTGTGAACCAAGCAGTCTCTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTCATATAAGT 651
 DB 94742 CTCCTGGAGCCAGCTGCTCTGCTGTGGAGGCGGAGTACTTAAGTCCCTTGTATTAAGT 94801
 QY 652 CAAAGTCGCAACAGAACCTCCCAAGGAAAGAAATTCAGGA---AAATGAATTTAGCTGT 708

DB 94802 CTGGAGCCAAAGAGCCAGCATTGGAGAGAGAGTGGCGAAGAGCCCGGAGAGTTGAGCTGT 94861
 QY 709 GAGGTATGTGGGCAGACATTTTAGAGTCGCTTTTGTATGTTGAGATCCACATGAGAACACAC 768
 DB 94862 GATGTGTGGGCAGACATTTCCAGTGGCTTTTGTATGTTGAGAGCCACATGAGAGCAT 94921
 QY 769 AAAGATTCTTTTCACTTACGGGTGTAACTGTGCGGAAGAAGATTCAAGAGAGCTTGTGTTT 828
 DB 94922 AAGGACTCTTTCACGTATGGGTGCAGATGTGCGGAGGAGATTCAAGGAGCGCTGGTTC 94981
 QY 829 CTTAAATATCATCTGGACRCATATATGSCAAATCGGGGCCAGAGCAAACTGCGAGCAA 888
 DB 94982 CTGAAGAACCATCTGGACACACAATGSCAAAGTCTGGCACCAGGAGCAAGCTTCAGCAA 95041
 QY 889 GGCTTGGAGAGTAGTCCAGCAACGATCAACAGAGTCTGTCAGGTGCACCGCGCCGAGAGC 948
 DB 95042 GGCATGGAG---AGTCCAGTCAACATCAATGAAGTGTCCAGCGCACGCCCTGGGAGC 95098
 QY 949 ATCTCTCTCTCTTTGCAAAATCTGATGTTTGTGCTTCTTATTTTCCAAATAAAGAAAGT 1008
 DB 95099 ATCTCCAGCCCTACAAGATCTGATGTTCTGCGGCTTCTCTTCCAAATAAGCAGAGC 95158
 QY 1009 CTAATTGAGCAGCCGACAGGTGCACACCAAACTGCTTTTCGTTACGACGAGCGCGCAG 1068
 DB 95159 CTCATTGAGCAGCAGCAAGGTTTCACGCCAAAGAACTGTCCCGAGTCCAGCAACGTTGCC 95218
 QY 1069 ACAGACTCTCCCAAGGAGGAATGCGTCTCGAGGAGGACTTCTCTGAGTGTGTTTCAAC 1128
 DB 95219 CTGATGACCCAGGAGAGACCCAGTCCCCGAGGAAAGATTGCTGCAAGTTTGAAC 95278
 QY 1129 TTGAGACCAAAATCTCACCTGAAACGGGGAAGAGCCCTGTGATGATGCAATCTCCAGCTC 1188
 DB 95279 TTGAGACCCAGATCACTGACAGTAGTACAGTGAAGCCCATGACCTGCATACCTCAGCTT 95338
 QY 1189 GATCCGTTCAACCTTCCAGGCTTGGCAKCTGGCTACCAAGAGAAAGTGTGCAATTTGC 1248
 DB 95339 GACCCGTTTCAACCTTACCCAGCATGGCAGTTGGCTTACCAAGAGAAAGTGGCCGTTGCC 95398
 QY 1249 C---AAGAGTGAAGGAATTTGGGCAAGAGGAGCAGCAGCAAGACGATTCGAGTTC 1305
 DB 95399 CAGGAAGAGGTGAAGAGTTCAGGCCAAGAGGAGCAGACATGACACTCATGCTCA 95458
 QY 1306 GAGAGGAGCTTGGAG----- 1321
 DB 95459 GAGAAAGAGGAAGTGGGGAATATGTTGGGGAATAGGCGGAAGGGTCTCGAAAGTCC 95518
 QY 1322 -AACAATAAGAACCATTTGTCAGGCTCTCGAAGAGAAAGAGAGTGCACACTCC 1380
 DB 95519 AAAAAAGTAAAGCAGTTGTCAGGTCTCTCCCAAGCAAAAGAGAGCCTTAGACATGCT 95578
 QY 1381 CAGCGCAAGCGCCCTCCGTGGACCGGATCCCAAGTTACCCAGTAGCAAGAGAGAGCC 1440
 DB 95579 AATAGTGAAGTGCTTCTGGGGATAGTACCCCAAGTTGTCAGTAGCAAGAGAGAGCC 95638
 QY 1441 ACTCACTGCTCGAGTGGGCAAGCTTTTCAAGAACTTACCAAGTGTGTTGCACTCC 1500
 DB 95639 AGGCACTGTTCTGAGTGCAGCAAGSCCTTCAGGACATACCACGCTCGTCTGCACTCG 95698
 QY 1501 AGGGTCC 1507
 DB 95699 AGGGTGC 95705

RESULT 15

BC072191
 LOCUS 3512 bp mRNA linear VRT 25-JUN-2004
 DEFINITION Xenopus laevis mitotic phosphoprotein 140, mRNA (cDNA clone
 MGC:80369 IMAGE:5074630), complete cds.
 ACCESSION BC072191
 VERSION BC072191.1 GI:47940271
 KEYWORDS MGC.
 SOURCE Xenopus laevis (African clawed frog)

Db 883 AATGCACCCCTCATGGAACATAGTAGATGCATTCCAAGGATTCTGGTTT 931

Search completed: June 29, 2005, 01:05:49
Job time : 4340.72 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 14:28:52 ; Search time 527.33 Seconds
(without alignments)
16917.381 Million cell updates/sec

Title: US-08-731-499-3

Perfect score: 1507
Sequence: 1 GCAGTTCTGGATTGACT.....GGTCTTGCACTCCAGGTC 1507

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1504.8	99.9	1507	2	Aav04698 Homo sapi
2	1387.2	92.1	36022	11	Acn44986 Human gen
3	1365	90.6	10282	2	Aav09023 Homo sapi
4	1337.2	88.7	5632	8	Acc50334 Breast ca
5	1337.2	88.7	5632	10	Add14635 Human src
6	1337.2	88.7	5632	13	Adr66761 Human pro
7	1337.2	88.7	5632	13	Adp65858 Human pro
8	1337.2	88.7	5632	12	Adp07267 Human znf
9	1132.6	75.2	5252	11	Acn44987 Human mrn
10	1087.2	72.1	3183	2	Aav09024 Homo sapi
11	573.6	38.1	26345	11	Acn44984 Mouse gen
12	537	35.6	3016	11	Acn44985 Mouse mrn
13	523.2	34.7	566	4	Aak81095 Human imm
14	324.2	21.5	14906	4	Aak81093 Human imm
15	215.6	14.3	530	12	Ach76516 Human gen
16	197.2	13.1	298	3	Aac07753 Human sec
17	156.4	10.4	469	2	Aav88042 EST clone
18	104.8	7.0	5994	13	Acn40355 Tumour-as
19	103.2	6.8	6033	13	Adr25650 Breast ca
20	103.2	6.8	8156	6	Abt07614 Human bre

c	21	97	6.4	251	12	ACH90216	Human gen
	22	70	4.6	70	13	ADR00072	ZNF217 PC
	23	70	4.6	2706	10	ACH07836	Novel cod
c	24	68.8	4.6	1014	12	ACH87715	Human gen
	25	68.8	4.6	1229	12	ACH86998	Human gen
c	26	68.8	4.6	2706	11	ADL71050	Gene enco
	27	68.8	4.6	2765	4	AAI58185	Human pol
	28	68.8	4.6	2765	5	ADQ98391	DNA enco
	29	68.8	4.6	2765	9	ADB48151	Novel hum
	30	68.8	4.6	3039	13	ADR25783	Breast ca
	31	68.8	4.6	3040	13	ADQ87590	Human tum
	32	68.8	4.6	3617	4	ABK42507	Genomic s
	33	68.8	4.6	3617	4	ABK42509	Genomic s
	34	68.8	4.6	3617	9	ADB60663	Connectiv
	35	68.8	4.6	3617	9	ADB60665	Connectiv
	36	63.6	4.2	2909	11	ADL71048	Gene enco
	37	59.2	3.9	508	12	ACH73808	Human gen
	38	59.2	3.9	772	4	AAS31069	Human dia
	39	59.2	3.9	2375	11	ADM02211	Human cdn
	40	59.2	3.9	10926	4	AAS65370	Human imm
	41	58.4	3.9	1995	4	AAS25920	Human cdn
	42	58.4	3.9	1995	8	ABX73261	Human nov
	43	53.6	3.6	1584	4	AAS26371	Human cdn
	44	53.6	3.6	1584	8	ABX73712	Human nov
c	45	53.6	3.6	2014	4	AAS25921	Human cdn

ALIGNMENTS

RESULT 1
AAV04698
ID AAV04698 standard; cdna; 1507 BP.
XX
AC AAV04698;
XX
DT 21-JUL-1998 (first entry)
XX
DE Homo sapiens 20q13 amplicon cc49 transcript.
XX
KW 20q13 amplicon; chromosome 20; tumour; detection; C2H2 zinc finger;
KW chromosomal abnormalities; probe; gene therapy; antisense inhibition;
KW treatment; age-related macular degeneration; retinitis pigmentata;
KW Leber's congenital anaurosis; ds.
XX
OS Homo sapiens.
XX
PN WO9802539-Al.
XX
PD 22-JAN-1998.
XX
PF 15-JUL-1997; 97WO-US012343.
XX
PR 15-JUL-1996; 96US-00680395.
PR 16-OCT-1996; 96US-00731499.
PR 17-JAN-1997; 97US-00785532.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Gray JW, Collins CC, Hwang S, Godfrey T, Kowbel D, Rommens J;
XX WPI; 1998-110587/10.
XX
PT New sequences from the 20q13 amplicon - used for detecting chromosomal
PT abnormalities, particularly tumours, and for developing products for
PT treating diseases.
XX
PS Claim 1; Page 61-62; 91pp; English.
XX
CC The sequence is that of a cdna sequence cc49, which was isolated from the
CC 20q13 amplicon. It shows homology to C2H2 zinc finger genes. It can be
CC used as a probe for the detection of chromosomal abnormalities at 20q13.
CC It and other sequences isolated from the 20q13 amplicon are consistently

CC amplified in primary tumours. These sequences are useful as probes or as
CC probe targets for monitoring the relative copy number of corresponding
CC sequences from a biological sample such as tumour cells. The sequences
CC can also be used in therapeutic applications for modulating the
CC expression of the endogenous gene or the activity of the gene product.
CC Examples of therapeutic approaches include antisense inhibition of gene
CC expression, gene therapy, and monoclonal antibodies that specifically
CC bind the gene products. The products can also be used in the treatment of
CC other diseases, e.g. age-related macular degeneration, Leber's congenital
CC amaurosis and retinitis pigmentation
XX

SQ Sequence 1507 BP; 445 A; 354 C; 364 G; 340 T; 0 U; 4 Other;

Query Match 99.9%; Score 1504.8; DB 2; Length 1507;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1507; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCAGGTTCGCGAATGACTCTTCTGCTCAATTAAGAACACTCATTTCAATGGAGACAAAGAG 60
Db 1 GCAGGTTCGCGAATGACTCTTCTGCTCAATTAAGAACACTCATTTCAATGGAGACAAAGAG 60
Qy 61 CACTAATGCTTTGCTGCTGATTCATATTTGAATCGAGGCAATGGGAAACCTGTATGCCCTTG 120
Db 61 CACTAATGCTTTGCTGCTGATTCATATTTGAATCGAGGCAATGGGAAACCTGTATGCCCTTG 120
Qy 121 TTTGTGGAAAGAACAGTGACACCATCACTGAGCTTCTTAAATAAATATTTGCTCTGGTTTGGAAACCCAGG 240
Db 121 TTTGTGGAAAGAACAGTGACACCATCACTGAGCTTCTTAAATAAATATTTGCTCTGGTTTGGAAACCCAGG 240
Qy 241 ACTGTTAGAGGTGAGTGACAGGCTTTACAGTGGCTTTAATCCAACTCCAGAAATGGCC 300
Db 241 ACTGTTAGAGGTGAGTGACAGGCTTTACAGTGGCTTTAATCCAACTCCAGAAATGGCC 300
Qy 301 AACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAACATGCCAACTCAATCCCT 360
Db 301 AACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAACATGCCAACTCAATCCCT 360
Qy 361 CTTAATGTACATGATGATGCCAAGAGTGATGGCAGCTCTCTTTGCCAGTCCGATGGAGATG 420
Db 361 CTTAATGTACATGATGATGCCAAGAGTGATGGCAGCTCTCTTTGCCAGTCCGATGGAGATG 420
Qy 421 GAGATGCCCTTGTCAATGAAGGGCCNCCTGTTGTCAATTCGAGCTACACAAAGAAAAA 480
Db 421 GAGATGCCCTTGTCAATGAAGGGCCNCCTGTTGTCAATTCGAGCTACACAAAGAAAAA 480
Qy 481 ATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCTGCGAGCCAGACCTTC 540
Db 481 ATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCTGCGAGCCAGACCTTC 540
Qy 541 ACACATTGAGAGCTTTAATAAATGCTTTAATGCAACACCGGCCCTACCTCTGTGAA 600
Db 541 ACACATTGAGAGCTTTAATAAATGCTTTAATGCAACACCGGCCCTACCTCTGTGAA 600
Qy 601 CCAGCAGTTCTTCGGGTTGAACGACAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGGGA 660
Db 601 CCAGCAGTTCTTCGGGTTGAACGACAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGGGA 660
Qy 661 ACAGAACCTCCCAAGAAAAAGAAATGCAAGGAAATGAATTTAGCTGTGAGTATGTGGG 720
Db 661 ACAGAACCTCCCAAGAAAAAGAAATGCAAGGAAATGAATTTAGCTGTGAGTATGTGGG 720
Qy 721 CAGACATTTAGAGTCCGCTTTTGGATGTTGAGATCCCATGAGAACACACAAAGATTTCTTC 780
Db 721 CAGACATTTAGAGTCCGCTTTTGGATGTTGAGATCCCATGAGAACACACAAAGATTTCTTC 780
Qy 781 ACTTACGGGTGTAACTGTGCGGAAGAAGATTCAGGAGCCCTTGGTTTCTTAAAAATCAC 840
Db 781 ACTTACGGGTGTAACTGTGCGGAAGAAGATTCAGGAGCCCTTGGTTTCTTAAAAATCAC 840

Qy 841 ATGCGACRCATATATGGCAAAATCGGGGCCAGAACTGCAGCAAGGCTTGAGAGT 900
Db 841 ATGCGACRCATATATGGCAAAATCGGGGCCAGAACTGCAGCAAGGCTTGAGAGT 900
Qy 901 AGTCCAGCAACGATCAACGAGGTGCTCCAGGTGACGCGGCGGAGAGCATCTCTCTCT 960
Db 901 AGTCCAGCAACGATCAACGAGGTGCTCCAGGTGACGCGGCGGAGAGCATCTCTCTCTCT 960
Qy 961 TGCATAATCTCATGTTGTTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTGAGCAC 1020
Db 961 TGCATAATCTCATGTTGTTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTGAGCAC 1020
Qy 1021 CGCAAGGTGCACACCAAAAAAATGCTTTTCGTATCCAGCAGCGCGCAGACACTCTCCA 1080
Db 1021 CGCAAGGTGCACACCAAAAAAATGCTTTTCGTATCCAGCAGCGCGCAGACACTCTCCA 1080
Qy 1081 CAAGGAGGAATGCCGTCTCCGAGGGAGGACTTCTGCAAGTGTTCATTTGAGACCAAAA 1140
Db 1081 CAAGGAGGAATGCCGTCTCCGAGGGAGGACTTCTGCAAGTGTTCATTTGAGACCAAAA 1140
Qy 1141 TCTCAACCTGAAACGGGGAAGAGCCTGTGAGATGCATCCCTCAGCTCGATCCGTTCAAC 1200
Db 1141 TCTCAACCTGAAACGGGGAAGAGCCTGTGAGATGCATCCCTCAGCTCGATCCGTTCAAC 1200
Qy 1201 ACCTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGTTCATTTGCAAGAACTGAAG 1260
Db 1201 ACCTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGTTCATTTGCAAGAACTGAAG 1260
Qy 1261 GAAATGGGCGCAAGAGGAGCACCGCAACGACGATTCGAGTTCGAGAAAGAGCTTGA 1320
Db 1261 GAAATGGGCGCAAGAGGAGCACCGCAACGACGATTCGAGTTCGAGAAAGAGCTTGA 1320
Qy 1321 GAAACAAATAGAACCAATTTGCGAGGCTCTCGCAGAGAAAGAGAGTGCACAACTCC 1380
Db 1321 GAAACAAATAGAACCAATTTGCGAGGCTCTCGCAGAGAAAGAGAGTGCACAACTCC 1380
Qy 1381 CACGCGAAGCGCCCTCCGTTGGAAGCGGATCCCAAGTTACCCAGTAGCAAGGAGAGGCC 1440
Db 1381 CACGCGAAGCGCCCTCCGTTGGAAGCGGATCCCAAGTTACCCAGTAGCAAGGAGAGGCC 1440
Qy 1441 ACTCATCTCTCCGAGTGGCGCAAGCTTTTCAGAACCTTACCACTCCAGCTGGTCTTGACATCC 1500
Db 1441 ACTCATCTCTCCGAGTGGCGCAAGCTTTTCAGAACCTTACCACTCCAGCTGGTCTTGACATCC 1500
Qy 1501 AGGGTCC 1507
Db 1501 AGGGTCC 1507
RESULT 2
ACN44986
ID ACN44986 standard; DNA; 36022 BP.
XX
AC ACN44986;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human genomic sequence hCG37127.
XX
KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003073826-A2.
XX
PD 12-SBP-2003.
XX
PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX

KW chromosomal abnormalities; probe; gene therapy; antisense inhibition;
 KW treatment; age-related macular degeneration; retinitis pigmentata;
 KW Leber's congenital amaurosis; zinc finger amplified in breast cancer; ds.
 XX
 OS Homo sapiens.
 XX
 PN W09802539-A1.
 XX
 XX 22-JAN-1998.
 PD
 XX 15-JUL-1997; 97WO-US012343.
 PF
 XX 15-JUL-1996; 96US-00680395.
 PR
 PR 16-OCT-1996; 96US-00731499.
 PR 17-JAN-1997; 97US-00785532.
 XX
 XX (REGC) UNIV CALIFORNIA.
 PA
 XX Gray JW, Collins CC, Hwang S, Godfrey T, Kowbel D, Rommens J;
 PI
 XX WPI; 1998-110587/10.
 DR
 DR P-PSDB; AAW23975.
 XX
 PT New sequences from the 20q13 amplicon - used for detecting chromosomal
 PT abnormalities, particularly tumours, and for developing products for
 PT treating diseases.
 XX
 PS Claim 1; Page 64-67; 91pp; English.
 XX
 CC The sequence is that of the genomic sequence of ZABC-1 (zinc finger
 CC amplified in breast cancer). It maps to the core of the 20q13.2 amplicon
 CC and is overexpressed in primary tumours and breast cancer cell lines
 CC having 20q13.2 amplification. The exact coding region for the genomic
 CC sequence is not given. The sequence can be used as a probe for the
 CC detection of chromosomal abnormalities at 20q13. It and other sequences
 CC isolated from the 20q13 amplicon are consistently amplified in primary
 CC tumours. These sequences are useful as probes or as probe targets for
 CC monitoring the relative copy number of corresponding sequences from a
 CC biological sample such as tumour cells. The sequences can also be used in
 CC therapeutic applications for modulating the expression of the endogenous
 CC gene or the activity of the gene product. Examples of therapeutic
 CC approaches include antisense inhibition of gene expression, gene therapy,
 CC and monoclonal antibodies that specifically bind the gene products. The
 CC products can also be used in the treatment of other diseases, e.g. age-
 CC related macular degeneration, Leber's congenital amaurosis and retinitis
 CC pigmentata
 XX
 SQ Sequence 10282 BP; 2820 A; 2222 C; 2191 G; 3047 T; 0 U; 2 Other;
 Query Match 90.6%; Score 1365; DB 2; Length 10282;
 Best Local Similarity 96.4%; Pred. No. 0;
 Matches 1455; Conservative 4; Mismatches 45; Indels 6; Gaps 6;
 QY 2 CAGGTTGCTGGATGACTCTTGTCTCAATTGAACACTATTCAATGAGACAAAGAGC 61
 DB |||||
 DB 2423 CAGGTTGCTCGAATTGACTCTTGTCTCAATTGAACACTATTCAATGAGACAAAGAGA 2482
 QY |||||
 QY 62 ACTAATGCTTTGTGCTGATTTCATATTGAATCGAGGCATTGGGAACCCCTGTATGCTTGT 121
 DB |||||
 DB 2483 ACTAATGCTTTGTGCTGCTTCATTTCATATTGAATCGAGGCATTGGGAACCCCTGTATGCTTGT 2542
 QY |||||
 QY 122 TTGTGGAAGAACCACTGACACCATCCTGAGCTTCTTAAAGATTCGAAGAATTAGAGG 181
 DB |||||
 DB 2543 TTGTGGAAGAACCACTGACACCATCCTGAGCTTCTTAAAGATTCGAAGAATTAGAGG 2602
 QY |||||
 QY 182 ACTATACACTTCTTTTGAACCTTTTATAATAATTTGCTCTGCTTTTGAACCCAGCA 241
 DB |||||
 DB 2603 ACTATACACTTCTTTTGAACCTTTTATAATAATTTGCTCTGCTTTTGAACCCAGG 2662
 QY |||||
 QY 242 CTGTTAGA-GGGTGAAGTCAAGGCTTAC-AGTGGCCTTAATCCAACTCCAGAAATGCC 299
 DB |||||
 DB 2663 CTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCCTTATTCCAACTCCAGAAATGCC 2722

QY 300 CAACGGAATTTGAGATTATATGCAATCGAAAGTCAGGAAACATGCCAACTCAATCCC 359
 DB |||||
 DB 2723 CAACGGAATTTGAGATTATATGCAATCGAAAGTCAGGAAACATGCCAACTCAATCCC 2782
 QY |||||
 DB 360 TCTTAATGTACATGGAT-GGCAAGAGTATTGGCAGCTCTCTTCCAGTCCGATGGAGA 418
 DB |||||
 DB 2783 TCTTAATGTACATGGATGGGCAGAGTATTGGCAGCTCTCTTGGCAGTCCGATGGAGA 2842
 QY |||||
 QY 419 TGGG-GATGCTTGTCAATGAAAGGCCCNCTGTTGTCAATTCGAGCTACACAAGAAA 477
 DB |||||
 DB 2843 TGGAGGATGCTTGTCAATGAAAGGCCGCTGTTGTTCCATTCGAGCTACACAAGAAA 2902
 QY |||||
 QY 478 AAAATGTCATCCGATCGAGGGGATATGCCCTTGGATTGATGATGCTTCTGAGCCAGACC 537
 DB |||||
 DB 2903 AAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTGGATTGATGCTTCTGAGCCAGACC 2961
 QY |||||
 DB 538 TTCAACATTCAGAAAGACCTTAAATAACATGCTCTTAATGCAACACCGGCTTACCCTCTGT 597
 DB |||||
 DB 2962 TTCAACATTCAGAAAGACCTTAAATAACATGCTCTTAATGCAACACCGGCTTACCCTCTGT 3021
 QY |||||
 QY 598 GAACGAGCAGTTCTTCCGGTTGAAGCAGAGTATCTCAGTCCGCTTGTATAAAAGTCAAGTG 657
 DB |||||
 DB 3022 GAACGAGCAGTTCTTCCGGTTGAAGCAGAGTATCTCAGTCCGCTTGTATAAAAGTCAAGTG 3081
 QY |||||
 QY 658 CGAACAGAACCTCCCAAGGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGTATGT 717
 DB |||||
 DB 3082 CGAACAGAACCTCCCAAGGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGTATGT 3141
 QY |||||
 QY 718 GGGCAGACATTTAGAGTCCGCTTTTGTGATGTTGAGATCCACATGAGAACACACAAAGATTCT 777
 DB |||||
 DB 3142 GGGCAGACATTTAGAGTCCGCTTTTGTGATGTTGAGATCCACATGAGAACACACAAAGATTCT 3201
 QY |||||
 DB 778 TTCACTTACGGGTGTAACATGTGCGGAAGAGATTCGAAGGAGCCTTGGTTTCTTAAAAAT 837
 DB |||||
 DB 3202 TTCACTTACGGGTGTAACATGTGCGGAAGAGATTCGAAGGAGCCTTGGTTTCTTAAAAAT 3261
 QY |||||
 QY 838 CACATCGGACRCATTAATGCGCAATTCGGGGCCAGAACCAACTGCACCAAGCCTTGGAG 897
 DB |||||
 DB 3262 CACATCGGACRCATTAATGCGCAATTCGGGGCCAGAACCAACTGCACCAAGCCTTGGAG 3321
 QY |||||
 DB 898 AGTAGTCCAGCAACGATCAACGAGGTCTCTCAGGTGACGCGGCCGAGAGCATCTCTCT 957
 DB |||||
 DB 3322 AGTAGTCCAGCAACGATCAACGAGGTCTCTCAGGTGACGCGGCCGAGAGCATCTCTCT 3381
 QY |||||
 QY 958 CTTTGCAAAATCTGATGCTTGTGCTTCTTCTAATTTCCAAATAAAGAAAGTCTAATTGAG 1017
 DB |||||
 DB 3382 CTTTGCAAAATCTGATGCTTGTGCTTCTTCTAATTTCCAAATAAAGAAAGTCTAATTGAG 3441
 QY |||||
 QY 1018 CACCGAAGTGCACACCAAAAACCTGCTTTTCGTTACCGAGCGCGCAGACACTCT 1077
 DB |||||
 DB 3442 CACCGAAGTGCACACCAAAAACCTGCTTTTCGTTACCGAGCGCGCAGACACTCT 3501
 QY |||||
 QY 1078 CCACAGGAGGAATGCCCTCTCGAGGGAGGACTTCTCTGAGTGTTCCTTCAACTTGAGACCA 1137
 DB |||||
 DB 3502 CCACAGGAGGAATGCCCTCTCGAGGGAGGACTTCTCTGAGTGTTCCTTCAACTTGAGACCA 3560
 QY |||||
 QY 1138 AAATCTCAACCTGAAAACGGGGAAAGAGCCTGTGATGATGATGCTTCCAGTCCGCTTC 1197
 DB |||||
 DB 3561 AAATCTCAACCTGAAAACGGGGAAAGAGCCTGTGATGATGATGCTTCCAGTCCGCTTC 3620
 QY |||||
 QY 1198 ACCACCTTCCAGGCTTGGCACTGGCTACCAAGAAAGTGGCTTTCCTTTCGCAAGAGTG 1257
 DB |||||
 DB 3621 ACCACCTTCCAGGCTTGGCACTGGCTACCAAGAAAGTGGCTTTCCTTTCGCAAGAGTG 3680
 QY |||||
 QY 1258 AAGGAATTTGGGGCAAGAGGGAGGACCCGACCAACGACGATTCGAGTTCGAGAGAGGAGCTT 1317
 DB |||||
 DB 3681 AAGGAATTTGGGGCAAGAGGGAGGACCCGACCAACGACGATTCGAGTTCGAGAGAGGAGCTT 3740
 QY |||||
 QY 1318 GGAGAAAACAAATAGAACCATTTGTGAGGCTCTCGAAGAGAGAAAGAGAGTGCACACAC 1377
 DB |||||
 DB 3741 GGAGAAAACAAATAGAGGAGTGTGTGAGGCTCTCGAAGAGACAAAGAGAGTGCACACAC 3800
 QY |||||
 QY 1378 TCCACGCGGAGGCGCCCTCTCGTGGACGCGGATCCCAAGTTTACCAGTAGCAAGGAGAG 1437

Db	3801	TCCACAGGCGAAGCGCCCTCGTGGACGCGGATCCCAAGTTACCCAGTAGCAAGGAGG	3860
Qy	1438	CCCACTCACTGCTCCGAGTGGCGCAAGCTTTTCAGAACTTACACCAAGCTGGTCTTGCAC	1497
Db	3861	CCCACTCACTGCTCCGAGTGGCGCAAGCTTTTCAGAACTTACACCAAGCTGGTCTTGCAC	3920
Qy	1498	TCCAGGTGCC 1507	
Db	3921	TCCAGGTGCC 3930	
RESULT 4			
ACC50334	ID	ACC50334 standard; cDNA; 5632 BP.	
XX	AC	ACC50334;	
XX	XX	12-JUN-2003 (first entry)	
XX	DB	Breast cancer associated cDNA sequence SEQ ID NO:505.	
XX	XX	Human; breast cancer; cytostatic; gene therapy; gene; ss.	
XX	OS	Homo sapiens.	
XX	XX	WO2003004989-A2.	
XX	PN	16-JAN-2003.	
XX	PD	21-JUN-2002; 2002WO-US019669.	
XX	PF	21-JUN-2001; 2001US-0299887P.	
XX	PR	27-JUN-2001; 2001US-0301572P.	
XX	PR	18-JUL-2001; 2001US-0306501P.	
XX	PR	25-SEP-2001; 2001US-0325002P.	
XX	PR	05-MAR-2002; 2002US-0362585P.	
XX	PR	14-MAY-2002; 2002US-0380039P.	
XX	PA	(MILL-) MILLENIUM PHARM INC.	
XX	PI	Lillie J, Gannavarapu M, Glatt K, Hoersh S, Kamatkar S;	
XX	PI	Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;	
XX	PI	Bast RC, Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;	
XX	XX	WPI: 2003-210381/20.	
XX	DR	P-PSDB; ABR47632.	
XX	PT	Breast cancer diagnosis or treatment by comparing the level of expression	
XX	PT	of a marker in a patient sample with that in the control non-breast	
XX	PT	cancer sample.	
XX	PS	Claim 1; SEQ ID NO 505; 128pp; English.	
XX	XX	The present invention describes a method for assessing whether a patient	
XX	CC	is afflicted with breast cancer. The method comprises comparing the level	
XX	CC	of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and	
XX	CC	ABR47386 to ABR47632) in a patient sample and the normal level of	
XX	CC	expression of the marker in a control non-breast cancer sample, where a	
XX	CC	significant increase in the level of expression of the marker in the	
XX	CC	patient sample and the normal level is an indication that the patient is	
XX	CC	afflicted with breast cancer. The breast cancer associated sequences from	
XX	CC	the present invention have cytostatic activities and can be used in gene	
XX	CC	therapy. The method is useful for diagnosing and treating breast cancer.	
XX	CC	N.B. the sequence data for this patent did not form part of the printed	
XX	CC	specification, but was obtained in electronic format directly from WIPO	
XX	CC	at ftp.wipo.int/pub/published_pct_sequences	
XX	SQ	Sequence 5632 BP; 1637 A; 1166 C; 1258 G; 1571 T; 0 U; 0 Other;	
Query Match 88.7%; Score 1337.2; DB 8; Length 5632;			
Best Local Similarity 97.2%; Pred. No. 0;			
Matches 1420; Conservative 3; Mismatches 32; Indels 6; Gaps 6;			

Query Match	Score	DB	Length
Best Local Similarity	97.2%		
Pred. No.	0		
Matches	1420		
Conservative	3		
Mismatches	32		
Indels	6		
Gaps	6		

QY 1127 ACTTGAGACCAAAATCTCACCTCTGAAACGGGGAAGCCTGTGCAGATGCATCCTCTCAGC 1186
 DB 1080 ACTTGAGACCAAAATCTCACCTCTGAAACGGGGAAGCCTGTGCAGATGCATCCTCTCAGC 1139
 QY 1187 TCGATCCGTTTCCACCACTTCCAGGCTTGGCAKCTGGCTTACCAAGGAAGTTCGCAATTT 1246
 DB 1140 TCGATCCGTTTCCACCACTTCCAGGCTTGGCAKCTGGCTTACCAAGGAAGTTCGCAATTT 1199
 QY 1247 GCCAAGAGTGAAGAAATTTGGGCAAGAGGAGCACCAGCAACGACGATTCGATTCGG 1306
 DB 1200 GCCAAGAGTGAAGAAATTTGGGCAAGAGGAGCACCAGCAACGACGATTCGATTCGG 1259
 QY 1307 AGAAGGAGCTTGGAGAAACAAATAGAACCACTTGTGCAGGCTCTCGCAAGAGAAGAGA 1366
 DB 1260 AGAAGGAGCTTGGAGAAACAAATAGAACCACTTGTGCAGGCTCTCGCAAGAGAAGAGA 1319
 QY 1367 AGTGCACACACTTCCACGCGCAAGCGCCCTCCGTGGACCGGATCCCAAGTTTACCAGTA 1426
 DB 1320 AGTGCACACACTTCCACGCGCAAGCGCCCTCCGTGGACCGGATCCCAAGTTTACCAGTA 1379
 QY 1427 GCAAGGAGAGCCCACTCACTGCTCCGAGTGGCGCAAGAGCTTTCAGAACCTTACCACGAGC 1486
 DB 1380 GCAAGGAGAGCCCACTCACTGCTCCGAGTGGCGCAAGAGCTTTCAGAACCTTACCACGAGC 1439
 QY 1487 TGGTCTTGCACCTCCAGGGTCC 1507
 DB 1440 TGGTCTTGCACCTCCAGGGTCC 1460

RESULT 5

ID ADD14635 standard; cDNA; 5632 BP.
 XX
 AC ADD14635;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE Human src biomarker polynucleotide SEQ ID NO:29.
 XX
 KW predictor set; protein tyrosine kinase activity modulator;
 KW protein tyrosine kinase pathway; protein tyrosine kinase; cytosolic;
 KW gene therapy; drug sensitivity; genetic profile; cancer; human; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2003062395-A2.
 XX
 PD 31-JUL-2003.
 XX
 PF 17-JAN-2003; 2003WO-US001981.
 XX
 PR 18-JAN-2002; 2002US-0350061P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Huang F, Fairchild CR, Lee FY, Shaw P;
 XX
 XX WPI; 2003-636735/60.
 DR P-PSDB; ADD14040.
 XX
 PT New polynucleotides and polypeptides for predicting the activity of
 PT compounds that interact with protein tyrosine kinases and/or protein
 PT tyrosine kinase pathways.
 XX
 PS Claim 2; SEQ ID NO 29; 139pp; English.
 XX
 CC The present invention describes a predictor set comprising a plurality of
 CC polynucleotides or polypeptides whose expression pattern is predictive of
 CC the response of cells to treatment with a compound that modulates protein
 CC tyrosine kinase activity or members of the protein tyrosine kinase
 CC pathway. Also described: (1) predicting whether a compound is capable of
 CC modulating the activity of cells, comprising obtaining a sample of cells,
 CC determining whether the cells express a plurality of markers, and

CC correlating the expression of the markers to the compound's ability to
 CC modulate the activity of the cells; (2) a plurality of cell lines for
 CC identifying polynucleotides and polypeptides whose expression levels
 CC correlate with compound sensitivity or resistance of cells associated
 CC with a disease state; and (3) identifying polynucleotides and
 CC polypeptides that predict compound sensitivity or resistance of cells
 CC associated with a disease state, comprising subjecting the plurality of
 CC cell lines to one or more compounds, analysing the expression pattern of
 CC a microarray of polynucleotides or polypeptides, and selecting
 CC polynucleotides or polypeptides that predict the sensitivity or
 CC resistance of cells associated with a disease state by using the
 CC expression pattern of the microarray. The polynucleotides and
 CC polypeptides have cytostatic activities, and can be used in gene therapy.
 CC The polynucleotides and polypeptides are useful in predicting the
 CC activity of compounds that interact with protein tyrosine kinases and/or
 CC protein tyrosine kinase pathways. These may be used in determining drug
 CC sensitivity in patients to allow the development of individualized
 CC genetic profiles which aid in treating diseases and disorders (e.g.
 CC cancer) based on patient response at a molecular level. The present
 CC sequence is used in the exemplification of the present invention.

XX Sequence 5632 BP; 1637 A; 1166 C; 1258 G; 1571 T; 0 U; 0 Other;

Query Match 88.7%; Score 1337.2; DB 10; Length 5632;
 Best Local Similarity 97.2%; Pred. No. 0;
 Matches 1420; Conservative 3; Mismatches 32; Indels 6; Gaps 6;

QY 52 GACAAAGAGCAGCTAATGCTTTGCTGATTCATATTTGAATCGAGGCAATGGGAACCTGT 111
 DB 1 GACAAAGAGAACTAATGCTTTGCTGATTCATATTTGAATCGAGGCAATGGGAACCTGT 60
 QY 112 TATGCTTTGTTGGGAAGAACAGTGACACCATCACTGAGCTTCCTAAAAGTTTCGAAG 171
 DB 61 TATGCTTTGTTGGGAAGAACAGTGACACCATCACTGAGCTTCCTAAAAGTTTCGAAG 120
 QY 172 AAGTTAGAGGACTATACACTTTCTTTTGAACCTTTTATAATAAATATTTGCTCTGG-TTTT 230
 DB 121 AAGTTAGAGGACTATACACTTTCTTTTGAACCTTTTATAATAAATATTTGCTCTGGTTT 180
 QY 231 GGAACCCAGGACTGTAGA-GGGTAGAGTGACAGGTCTTAC-AGTGGCCCTTAATCCAATC 288
 DB 181 GGAACCCAGGCTGTAGAGGGGTGAGTGACAAGTCTTTACAAGTGGCCCTTAATCCAATC 240
 QY 289 CAGAAATTTGCCCAACGGAACCTTTGAGATTATATGCAATCGAAAGTGACAGAAACATGCC 348
 DB 241 CAGAAATTTGCCCAACGGAACCTTTGAGATTATATGCAATCGAAAGTGACAGAAACATGCC 300
 QY 349 AACTCAATCCCTCTTAATGTACATGGAT-GGCCAAGAGTGAATGGCAGCTCTCTTGCAG 407
 DB 301 AACTCAATCCCTCTTAATGTACATGGATGGCCAGAGTGAATGGCAGCTCTCTTGCAG 360
 QY 408 TCCGATGGAGATGGA-GATGCTTGTCAATGAAAGGCCCTCTGTGTCAATTCGAGCT 466
 DB 361 TCCGATGGAGATGGAGGATGCTTGTCAATGAAAGGCCCTCTGTGTTCATTCGAGC 420
 QY 467 ACACAAGAAAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCT 526
 DB 421 TACACAGAAAAAATGTCT-ATCCAATCGAGGGGTATATGCCCTTGGATTGCATGTTCT 479
 QY 527 GCAGCCAGACCTTCCACATTCAGAAGACCTTAATAAATCATGTCTTAAATGCAACACCGGC 586
 DB 480 GCAGCCAGACCTTCCACATTCAGAAGACCTTAATAAATCATGTCTTAAATGCAACACCGGC 539
 QY 587 CTACCTCTGTGAACCAAGCAGTCTTCCGGTGTGAAGCAGAGTATCTCAGTCCGCTTGATA 646
 DB 540 CTACCTCTGTGAACCAAGCAGTCTTCCGGTGTGAAGCAGAGTATCTCAGTCCGCTTGATA 599
 QY 647 AAAGTCAAGTCCGAACAGAACTCCCAAGGAAAAAGAAATTCGAAGAAATGAATTAGCT 706
 DB 600 AAAGTCAAGTCCGAACAGAACTCCCAAGGAAAAAGAAATTCGAAGAAATGAATTAGCT 659
 QY 707 GTGAGGTATGTGGGAGACATTTAGAGTGCCTTTTGTGTTGAGATCCATGAGAACAC 766

QY 231 GGAACCCAGGACTGTTAGA-GGGTCAGTGACAGGCTTTAC-AGTGGCCTTAATCCAACTC 288
DB |||||||
QY 181 GGAACCCAGGCTGTTAGAGGGGTGAGTGACAAGTCTTTACAAGTGGCCTTATTTCCAATC 240
DB |||||||
QY 289 CAGAAATTTGCCCAACGGAATTTTGAATATATATGCAATCGAAAGTGACAGGAAAATGCC 348
DB |||||||
QY 241 CAGAAATTTGCCCAACGGAATTTTGAATATATATGCAATCGAAAGTGACAGGAAAATGCC 300
QY 349 AACTCAATCCCTCTTAATGTATACATGGAT-GGCCAAGAGTGATGGCAGCTCTCTTGGCAG 407
DB |||||||
QY 301 AACTCAATCCCTCTTAATGTATACATGGATGGGCGCAGAAAGTGATGGCAGCTCTCTTGGCAG 360
QY 408 TCCGATGAGATGGA-GATGCCTCTGCAATGAAGGGCCNCTGTGTCAATTCGAGCT 466
DB |||||||
QY 361 TCCGATGAGATGGAAGGATGCTTGTCAATGAAGGGACCGCTGTGTTCANTCCGAGC 420
QY 467 ACACAAGAAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCT 526
DB |||||||
QY 421 TACAACAAGAAAAATGTC-ATCCAATCGAGGGGTATATGCCCTTGGATTGCATGTTCT 479
QY 527 GCAGCAGACCTTCACACATTCAGAAGACCTTAATAAATGATGCTTAATGCAACACCGGC 586
DB |||||||
QY 480 GCAGCAGACCTTCACACATTCAGAAGACCTTAATAAATGATGCTTAATGCAACACCGGC 539
QY 587 CTACCTCTGTGAACACGAGTCTTCGSGTTGAAGCAGAGTATCTCAGTCGCTTGATA 646
DB |||||||
QY 540 CTACCTCTGTGAACACGAGTCTTCGSGTTGAAGCAGAGTATCTCAGTCGCTTGATA 599
QY 647 AAAGTCAAGTGCAGAACAGACCTCCCAAGGAAAAGAAATTGCAAGGAAAATGAATTTAGCT 706
DB |||||||
QY 600 AAAGTCAAGTGCAGAACAGACCTCCCAAGGAAAAGAAATTGCAAGGAAAATGAATTTAGCT 659
QY 707 GTGAGGTATGGGCGAGACATTTAGATGCGCTTTTGTGATGTTGAGATCCACATGAGAACAC 766
DB |||||||
QY 660 GTGAGGTATGGGCGAGACATTTAGATGCGCTTTTGTGATGTTGAGATCCACATGAGAACAC 719
QY 767 ACAAGATTTCTTCACTTACGGGTGTAACATGTGCGGAAGAGATTCAAGGAGCCTTCGT 826
DB |||||||
QY 720 ACAAGATTTCTTCACTTACGGGTGTAACATGTGCGGAAGAGATTCAAGGAGCCTTCGT 779
QY 827 TTCTTAAAAATCACATGCGGACRCATAATGGCAAAATCGGGGCCACAGAACCAATGCGAGC 886
DB |||||||
QY 780 TTCTTAAAAATCACATGCGGACACATAATGGCAAAATCGGGGCCACAGAACCAATGCGAGC 839
QY 887 AAGGCTTGAGAGTAGTCCAGCAACGATCAACGAGGTGTCAGTGTCACGCGGCGGAGA 946
DB |||||||
QY 840 AAGGCTTGAGAGTAGTCCAGCAACGATCAACGAGGTGTCAGTGTCACGCGGCGGAGA 899
QY 947 GCATCTCTCTCTTGCAGAAATCTCATGTTTGTGGCTTCTTATTTCCAAATAAGAAA 1006
DB |||||||
QY 900 GCATCTCTCTCTTGCAGAAATCTCATGTTTGTGGCTTCTTATTTCCAAATAAGAAA 959
QY 1007 GTCTAATTTGAGCAGCGCAAGGTGACACCAAAAAAATGCTTTTCGGTACCAGCAGCGCG 1066
DB |||||||
QY 960 GTCTAATTTGAGCAGCGCAAGGTGACACCAAAAAAATGCTTTTCGGTACCAGCAGCGCG 1019
QY 1067 AGACAGACTCTCCACAGAGGAATGCGGTCTCTCAGGAGGAGACTCTCTGCAATGTTTCA 1126
DB |||||||
QY 1020 AGACAGACTCTCCACAGAGGAATGCGGTCTCTCAGGAGGAGGACTCTCTGCAATGTTTCA 1079
QY 1127 ACTTGAGACCAAAATCTCACCTGAAACGGGGAAGAGCTGTTCAGATGCAATCCCTCAGC 1186
DB |||||||
QY 1080 ACTTGAGACCAAAATCTCACCTGAAACGGGGAAGAGCTGTTCAGATGCAATCCCTCAGC 1139
QY 1187 TCGATCCGTTTCAACACCTTTCAGAGTTGGCAKCTGGCTACCAAGGAAGAGTTGCCATTT 1246
DB |||||||
QY 1140 TCGATCCGTTTCAACACCTTTCAGAGTTGGCAKCTGGCTACCAAGGAAGAGTTGCCATTT 1199
QY 1247 GCCAAGAGTAGGAAGATTTGGGCAAGAGGAGCACCAGCAACGACATTCGATTTCG 1306
DB |||||||
QY 1200 GCCAAGAGTAGGAAGATTTGGGCAAGAGGAGCACCAGCAACGACATTCGATTTCG 1259
QY 1307 AGAAGGAGCTTGGAGAAACAAATAAGAACCATTTGTGAGGCGCTCTCGCAAGAGAAAGAGA 1366

DB 1260 AGAAGGAGCTTGGAGAAACAATAAGGCAAGTGTGCAAGGCTCTCGCAAGAAAGAGA 1319
QY 1367 AGTGCAAAACACTCCCAACGCGAAGCGCCCTCCGTGGACGCGATCCCAAGTTTACCCAGTA 1426
DB 1320 AGTGCAAAACACTCCCAACGCGAAGCGCCCTCCGTGGACGCGATCCCAAGTTTACCCAGTA 1379
QY 1427 GCAAGGAGAACCCCACTCACTGCTCCGAGTGGGCAAGGCTTCAGAACCTTACCACCGC 1486
DB 1380 GCAAGGAGAACCCCACTCACTGCTCCGAGTGGGCAAGGCTTCAGAACCTTACCACCGC 1439
QY 1487 TGGTCTTCACTCCAGGGTCC 1507
DB 1440 TGGTCTTCACTCCAGGGTCC 1460

RESULT 7
ADR65858
ID ADR65858 standard; DNA; 5632 BP.
XX ADR65858;
XX
DT 02-DEC-2004 (first entry)
XX Human prostatic carcinoma derived DNA SEQ ID 54 #1.
XX human; cytostatic; diagnosis; prostatic cancer;
KW differential expression analysis; ds.
XX Homo sapiens.
XX WO2004076614-A2.
XX PD 10-SEP-2004.
XX
PF 22-FEB-2004; 2004WO-DE000433.
PR 27-FEB-2003; 2003DE-01009985.
PR 14-MAY-2003; 2003DE-01022134.
XX
PA (HINZ/) HINZMANN B.
PA (DAHL/) DAHL E.
PA (ROSE/) ROSENTHAL A.
PA (HERM/) HERMANN K.
PA (PILA/) PILARSKY C.
XX
PI Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarsky C, Specht T;
PI Schmitt A, Beckmann G, Bruemendorf T, Kinnemann H, Roepcke S;
PI Xinzhang L, Staub E;
XX
DR WPI; 2004-653386/63.
XX
PT New nucleic acids, and encoded proteins, from prostatic cancer tissue,
PT useful for diagnosis, treatment and in screening for specific binding
PT agents.
XX
PS Claim 1; Page 253-254; 1607pp; German.
XX
CC This invention describes novel cytostatic polynucleotide and polypeptide
CC sequences which can be used in a method for diagnosing prostatic cancer
CC or the risk of developing prostatic cancer. Diagnosis is based on
CC determining over transcription or over expression of the sequences in
CC prostatic tissue. Screening for inhibitors of the sequences or detection
CC substances involves a binding assay, any compounds that bind are
CC selected, optionally after deconvolution of mixtures. Detection of a
CC predetermined minimum level of the reporter indicates the presence of
CC tumour cells. Inhibitors can be chosen from antisense oligonucleotides,
CC short-interfering RNA or ribozymes; an organic molecule of molecular
CC weight below 5000, preferably 300, that binds to the polypeptide; an
CC aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the
CC polypeptide, preferably humanised or human; an anti-idiotypic, non-human
CC (monoclonal) antibody directed against Ab or any of the above derivatised
CC with a reporter group, cell toxin, immunostimulatory molecules and/or

CC radioisotope. The polynucleotides are identified in human prostatic
CC cancer by differential expression analysis, using DNA microarrays,
CC between normal and tumorous tissues, with (over)expression being detected
CC by quantitative PCR. Analysis of prostatic cancer samples showed that
CC CD24 was upregulated in many of them. Sections of tissue, isolated from
CC prostatic cancer patients, or subjects at risk, were incubated
CC sequentially with anti-human CD4 murine monoclonal antibodies;
CC biotinylated second antibody; streptavidin-conjugated horseradish
CC peroxidase and then diaminobenzidine as colour former (brown). The
CC samples were counterstained with hemalum (blue). Malignant cells stained
CC strongly but non-malignant cells only weakly. In 15 of 63 samples of
CC adenocarcinoma, membrane and cytoplasmic staining was very strong, and
CC lymph node metastases were also stained. ADR65805-ADR66954 represent the
CC polynucleotide and polypeptide sequences used in the method of the
CC invention.

XX
SQ Sequence 5632 BP; 1637 A; 1166 C; 1258 G; 1571 T; 0 U; 0 Other;

Query Match 88.7%; Score 1337.2; DB 13; Length 5632;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1420; Conservative 3; Mismatches 32; Indels 6; Gaps 6;

QY 52 GACAAAGACCTAATGCTTGTGCTGATTCATATTTGAATCGAGCATTTGGACCCCTG 111
DB 1 GACAAAGAACCTAATGCTTGTGCTGATTCATATTTGAATCGAGCATTTGGACCCCTG 60
QY 112 TATGCCCTTGTGTGAAAGAACCCAGTGACACCATCAGCTTCTCTAAAGTTTCGAAG 171
DB 61 TATGCTTGTGTGAAAGAACCCAGTGACACCATCAGCTTCTCTAAAGTTTCGAAG 120
QY 172 AAGTTAGAGGACTATACATCTTCTTTTGAATTTTATATAATATTTGCTCTGG-TTTT 230
DB 121 AAGTTAGAGGACTATACATCTTCTTTTGAATTTTATATAATATTTGCTCTGGTTTT 180
QY 231 GGAACCCAGGACTGTAGG-GGGTAGTGACAGGCTTTAC-AGTGGCCTTAATCCAACTC 288
DB 181 GGAACCCAGGCTGTGTAGAGGGGTAGTGACAGGCTTTACAGTGGCCCTTATTCCAACTC 240
QY 289 CAGAAATTCGCCCAACGGAATTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 348
DB 241 CAGAAATTCGCCCAACGGAATTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300
QY 349 AACTCAATCCCTCTTAATGTATCATGGAT-GGCCAAGAGTGATTTGGCAGCTCTCTTGGCCAG 407
DB 301 AACTCAATCCCTCTTAATGTATCATGGATGGCCAGAGTGATTTGGCAGCTCTCTTGGCCAG 360
QY 408 TCCGATGGAGATGGA-GATGCCCTGTCTAATGAAGGCCCNCTGTGTCAATTCGAGCT 466
DB 361 TCCGATGGAGATGGAGGATGGCTTGTCAATGAAGGGACCGCTGTGTTCATTCGAGC 420
QY 467 ACACAAAGAAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCT 526
DB 421 TACACAGAAAAAATGTCT-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCT 479
QY 527 GCAGCCAGACCTTTCACATTCAGAAAGCCTTAAATAACATGTCTTTAATGCAACACCCGCG 586
DB 480 GCAGCCAGACCTTTCACATTCAGAAAGCCTTAAATAACATGTCTTTAATGCAACACCCGCG 539
QY 587 CTACCTCTGTGAACCCAGCAGTCTTTCGGGTGGAAGCAGAGTATCTCAGTCCGCTTGATA 646
DB 540 CTACCTCTGTGAACCCAGCAGTCTTTCGGGTGGAAGCAGAGTATCTCAGTCCGCTTGATA 599
QY 647 AAAGTCAAGTCGGAACAGAACCTCCCAAGGAAAAAGATTGCAAGAAATGAATTTAGCT 706
DB 600 AAAGTCAAGTCGGAACAGAACCTCCCAAGGAAAAAGATTGCAAGAAATGAATTTAGCT 659
QY 707 GTGAGGTATGTGGGACGACATTTAGAGTCGCTTTTGTATTTGAGATFCCACATGAGAACAC 766
DB 660 GTGAGGTATGTGGGACGACATTTAGAGTCGCTTTTGTATTTGAGATFCCACATGAGAACAC 719
QY 767 ACAAGATTTCTTTCACTTACGGGTGTAAATGTGCGGAAGAGATTCAAGGAGCCTTGGT 826
DB 720 ACAAGATTTCTTTCACTTACGGGTGTAAATGTGCGGAAGAGATTCAAGGAGCCTTGGT 779

QY 827 TTCTTTAAAAATCACATGCGGACRCATATGCAAAATCGGGGCCAGAAACAACCTGCAGC 886
DB 780 TTCTTTAAAAATCACATGCGGACACATATGCAAAATCGGGGCCAGAAACAACCTGCAGC 839
QY 887 AAGGCTTGGAGAGTAGTCCAGCAACATCAACAGAGTCTCCAGGTGACGCGGCCGAGA 946
DB 840 AAGGCTTGGAGAGTAGTCCAGCAACATCAACAGAGTCTCCAGGTGACGCGGCCGAGA 899
QY 947 GCATCTCTCTCTCTTTCGCAAAATCTCATGTTTGTGGCTTCTTATTTCCAAATAAAGAA 1006
DB 900 GCATCTCTCTCTCTTACAAAATCTGCATGTTTGTGGCTTCTTATTTCCAAATAAAGAA 959
QY 1007 GTCTAATTTGAGCACCGCAGAGTGTCACACCAAAAAAATCTGCTTTTCGGTACACGAGCGCGC 1066
DB 960 GTCTAATTTGAGCACCGCAGAGTGTCACACCAAAAAAATCTGCTTTTCGGTACACGAGCGCGC 1019
QY 1067 AGACAGACTCTCCACAGGAGGAATGCGTCTCTCGAGGGAGGACTTCTCTGCAAGTTGTTCA 1126
DB 1020 AGACAGACTCTCCACAGGAGGAATGCGTCTCTCGAGGGAGGACTTCTCTGCAAGTTGTTCA 1079
QY 1127 ACTTGAGACCAAAATCTCACCCCTGAAACGGGGAAGAGCCTGTGATGCATCCCTCAGC 1186
DB 1080 ACTTGAGACCAAAATCTCACCCCTGAAACGGGGAAGAGCCTGTGATGCATCCCTCAGC 1139
QY 1187 TCGATCCGTTTCAACCACTTTCAGAGGCTTGGCACTCTGCTACCAAGGAAGTTGCCATTT 1246
DB 1140 TCGATCCGTTTCAACCACTTTCAGAGGCTTGGCACTCTGCTACCAAGGAAGTTGCCATTT 1199
QY 1247 GCCAAGAGAGTGAAGAAATTTGGGGCAAGAGGAGGACACCGAACGAGATTGAGTTCCG 1306
DB 1200 GCCAAGAGAGTGAAGAAATCGGGGCAAGAGGAGGACACCGAACGAGATTGAGTTCCG 1259
QY 1307 AGAAGGAGCTTGGAGAAACAAATAGAACCATTTGTGAGGCGCTCTCGCAAGAGAAAGAGA 1366
DB 1260 AGAAGGAGCTTGGAGAAACAAATAGAACCATTTGTGAGGCGCTCTCGCAAGAGAAAGAGA 1319
QY 1367 AGTGAACACACTCCCAACGCGGAGGCGCTCTCGTGGAGCGCGATCCCAAGTTTACCCAGTA 1426
DB 1320 AGTGAACACACTCCCAACGCGGAGGCGCTCTCGTGGAGCGCGATCCCAAGTTTACCCAGTA 1379
QY 1427 GCAAGGAGAGCCCACTCACTGCTCCGAGTGCGGCAAGAGCTTTTCAGAACTTACCACAGC 1486
DB 1380 GCAAGGAGAGCCCACTCACTGCTCCGAGTGCGGCAAGAGCTTTTCAGAACTTACCACAGC 1439
QY 1487 TGGTCTTGCACCTCCAGGGTCC 1507
DB 1440 TGGTCTTGCACCTCCAGGGTCC 1460

RESULT 8

ADP07267
ID ADP07267 standard; DNA; 5653 BP.

XX AC ADP07267;

XX DT 29-JUL-2004 (first entry)

XX DE Human ZNF217 DNA.

XX KW ds; proliferative disease; breast; methylation; CpG; bisulfite; human.

XX OS Homo sapiens.

XX PN DE10255104-A1.

XX PD 11-MAR-2004.

XX PF 26-NOV-2002; 2002DE-01055104.

XX PR 27-AUG-2002; 2002DE-01039313.

XX PA (EPIC-) EPIGENOMICS AG.

XX
PI Mater S;
XX
DR WPI; 2004-284340/27.
DR GENBANK; NM_006526.
XX
PT Analyzing proliferative diseases of breast cells, useful e.g. for
PT diagnosis, prognosis and treatment of breast cancer, by determining
PT methylation status of specific genes.
XX
PS Claim 1; Page; 22pp; German.
XX
CC This invention describes a novel method of analysing proliferative
CC diseases of breast cells by determining the methylation status of certain
CC genes. The invention also describes nucleic acids or their complements,
CC oligomers, especially oligonucleotides or peptide nucleic acid oligomers,
CC that hybridise to, or are identical with, any of the nucleic acids, the
CC preparation of an oligomer array for analysing proliferative diseases of
CC breast cells that are associated with the methylation status of CpG
CC dinucleotides of the genes by bonding at least one oligomer to a solid
CC phase and a kit comprising a bisulfite reagent and the oligomer. The
CC nucleic acids are genomic sequences (5' - and/or regulatory and/or CpG-
CC rich regions). The base sequence of the oligomer includes at least one
CC CpG island, especially with C in the middle third of the sequence. The
CC process involves a genomic DNA sample treated chemically, specifically
CC with a bisulfite reagent, to convert non-methylated C to uracil or some
CC other base having base-pairing properties different from those of C.
CC Fragments of the treated DNA are then amplified, using primers and a
CC polymerase and the methylation status of the genomic CpG dinucleotides is
CC determined by analysis of the amplicons, particularly by hybridisation to
CC the oligomer, optionally with extension of the hybridised oligomer by at
CC least one base, or detection is by sequencing. The amplification may use
CC methylation-specific primers. Alternatively, genomic DNA is extracted
CC from a sample and digested with methylation-specific restriction enzymes,
CC then the digestion fragments detected, optionally after amplification. In
CC either method, more than 10 fragments of 100-200 bp are amplified in a
CC single reaction vessel, using a heat-resistant DNA polymerase in PCR. The
CC amplicons carry detectable markers, e.g. fluorophores, radioisotopes
CC and/or releasable fragments of known mass that can be detected by mass
CC spectrometry. The method is used for characterisation, classification,
CC differentiation, staging, phase-estimation, diagnosis and/or therapy of
CC proliferative diseases of breast cells. The method provides very specific
CC classification of proliferative diseases, allowing better treatment. It
CC can both characterise methylation status and detect single-nucleotide
CC polymorphisms. This sequence represents human gene used to illustrate the
CC method of the invention. NOTE: This sequence does not appear in the
CC printed specification but has been retrieved from Genbank.
XX
SQ Sequence 5653 BP; 1656 A; 1167 C; 1258 G; 1572 T; 0 U; 0 Other;

Query Match 88.7%; Score 1337.2; DB 12; Length 5653;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1420; Conservative 3; Mismatches 32; Indels 6; Gaps 6;
QY 52 GACAAAGAGCACTAATGCTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCCTG 111
DB 1 GACAAGAGAACTAATGCTTGTGCTGATTCATATTTGAATCGAGGCATTGGGAACCCCTG 60
QY 112 TATGCTTGTGTTGGGAAGAACCAAGTACACCATCCTGAGCTTCCCTAAAAGTTTCGAAG 171
DB 61 TATGCTTGTGTTGGGAAGAACCAAGTACACCATCCTGAGCTTCCCTAAAAGTTTCGAAG 120
QY 172 AAGTTAGAGGACTATACACTTCTTTTGAACCTTTTATATAATAATTTGCTCTGG-TTTT 230
DB 121 AAGTTAGAGGACTATACACTTCTTTTGAACCTTTTATATAATAATTTGCTCTGGTTTTT 180
QY 231 GGAAACCCAGGACTGTTAGA-GGGTCAGTGACAGGTCCTTAC-AGTGCCTTTAATCAACTC 288
DB 181 GGAAACCCAGGCTGTTAGAGGGGTGAGTGACAAGTCTTACAAGTGGCCCTTATTCCACTC 240
QY 289 CAGAAATTTGCCCAACGGAACTTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 348
DB 241 CAGAAATTTGCCCAACGGAACTTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300

QY 349 AACTCAATCCCTCTTAATGTATACATGGAT-GGCCAAGAGTGATGGCAGCTCTCTTGGCAG 407
DB 301 AACTCAATCCCTCTTAATGTATACATGGATGGGCGAAGAGTGATGGCAGCTCTCTTGGCAG 360
QY 408 TCCGATGAGATGGA-GATGCTTGTCAATGAAAGGGCCNCTGTGTGTCATATTCGAGCT 466
DB 361 TCCGATGAGATGAGGATGCTTGTCAATGAAAGGGACCGCTGTGTTCATTCGAGC 420
QY 467 ACACAAAGAAAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGTGATTCGATGTTCT 526
DB 421 TACACAGAAAAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTGGATTGATGTTCT 479
QY 527 GCAGCCAGACCTTCAACATTCAGAGACCTTATTAACATCTCTTAATGAAACACCGC 586
DB 480 GCAGCCAGACCTTCAACATTCAGAGACCTTATTAACATCTCTTAATGAAACACCGC 539
QY 587 CTACCTCTGTGAACCAAGCAGTCTTTCGGGTGGAAGCAGAGTATCTCAGTCCGCTTGATA 646
DB 540 CTACCTCTGTGAACCAAGCAGTCTTTCGGGTGGAAGCAGAGTATCTCAGTCCGCTTGATA 599
QY 647 AAAGTCAAGTGCAGAACAACTCCCAAGAAAAAATTCGAAGAAAAATGAATTTAGCT 706
DB 600 AAAGTCAAGTGCAGAACAACTCCCAAGAAAAAATTCGAAGAAAAATGAATTTAGCT 659
QY 707 GTGAGTATGTGGGAGACATTTAGAGTCGCTTTTGTGATTTGAGATCCACATGAAACAC 766
DB 660 GTGAGTATGTGGGAGACATTTAGAGTCGCTTTTGTGATTTGAGATCCACATGAAACAC 719
QY 767 ACAAGATTTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAAGATTCAAGAGCCTTGGT 826
DB 720 ACAAGATTTCTTTCACTTACGGGTGTAACATGTGCGGAAGAAAGATTCAAGAGCCTTGGT 779
QY 827 TTCTTAAAAATCAATGCGGACRCAATATGGAATCGGGGCCAAGAACAAATTCGAGC 886
DB 780 TTCTTAAAAATCAATGCGGACACATAATGGCAATATGGCAATATGGCAATATGGCAAT 839
QY 887 AAGCTTGGAGAGTAGTCCAGCACGATCAACAGAGTCTCCAGGTGACCGCGGCCGAGA 946
DB 840 AAGCTTGGAGAGTAGTCCAGCACGATCAACAGAGTCTCCAGGTGACCGCGGCCGAGA 899
QY 947 GCATCTCTCTCTTTCGCAAAATCTGCATGTTGTGGCTTCTCTATTTCCAAATAAGAAA 1006
DB 900 GCATCTCTCTCTTTCGCAAAATCTGCATGTTGTGGCTTCTCTATTTCCAAATAAGAAA 959
QY 1007 GTCTAATTTGAGACCGCAAGGTGACACAAAAAATCTGTTTGGTACCGACAGCGCGC 1066
DB 960 GTCTAATTTGAGACCGCAAGGTGACACAAAAAATCTGTTTGGTACCGACAGCGCGC 1019
QY 1067 AGACAGCTCTCCCAAGGAGGAATGCGCTCTCGAGGGAGGACCTTCTGCAAGTTGTTCA 1126
DB 1020 AGACAGCTCTCCCAAGGAGGAATGCGCTCTCGAGGGAGGACCTTCTGCAAGTTGTTCA 1079
QY 1127 ACTTGAGACCAAAATCTCACCTCGAAACGGGGAAGAACCTGTGATGATGCAATCCCTCAGC 1186
DB 1080 ACTTGAGACCAAAATCTCACCTCGAAACGGGGAAGAACCTGTGATGATGCAATCCCTCAGC 1139
QY 1187 TCGATTCCTTCCACCTTTCAGGGCTTGGCACTGGCTTACCAAGGAAGTTGTCATTT 1246
DB 1140 TCGATTCCTTCCACCTTTCAGGGCTTGGCACTGGCTTACCAAGGAAGTTGTCATTT 1199
QY 1247 GCCAAGAGTCAAGGAATTTGGGCAAGAGGAGGACCGAACCAACCGATTCGAGTTCCG 1306
DB 1200 GCCAAGAGTCAAGGAATTTGGGCAAGAGGAGGAGGACCGAACCAACCGATTCGAGTTCCG 1259
QY 1307 AGAAGAGCTTTGGGAAAAAATAATTAAGAACCAATTTGTCAGGCTCTTCGAAAGAAAGAGA 1366
DB 1260 AGAAGAGCTTTGGGAAAAAATAATTAAGAGGAGTTTGTGAGGCTCTTCGCAAGAAAGAGA 1319
QY 1367 AGTGCAAAACATCCCAACCGGCAAGGCGCTTGGTGAGCGCGATCCCAAGTTACCCAGTA 1426
DB 1320 AGTGCAAAACATCCCAACCGGCAAGGCGCTTGGTGAGCGCGATCCCAAGTTACCCAGTA 1379

Qy	1427	GCAAGGAGAGAGCCCACTCACTGCTCCGAGTGGCGGAAAGCTTTTCAGAACTTACCAACCAGC	1487
Db	1380	GCAAGGAGAGAGCCCACTCACTGCTCCGAGTGGCGGAAAGCTTTTCAGAACTTACCAACCAGC	1439
Qy	1487	TGCTCTTGGCACTCCAGGGTCC	1507
Db	1440	TGGTCTTGGCACTCCAGGGTCC	1460
RESULT 9			
ACN44987			
ID	ACN44987	standard; cDNA; 5252 BP.	
XX	AC		
XX	ACN44987;		
XX			
DT	18-NOV-2004	(first entry)	
XX			
DE	Human mRNA sequence	hCT1950762.	
XX			
XX	Cytostatic; carcinoma; lymphoma; cancer; human; gene; ds.		
KW	Homo sapiens.		
OS			
XX			
PN	W02003073826-A2.		
XX			
PD	12-SEP-2003.		
XX			
XX	28-FEB-2003; 2003WO-US006235.		
PF			
XX			
PR	01-MAR-2002; 2002US-00087192.		
XX			
PA	(SAGR-) SAGRES DISCOVERY.		
XX			
PI	Morris DW;		
XX			
XX	WPI; 2003-328604/31.		
DR			
XX			
PT	Recombinant nucleic acid useful for diagnosis and treatment of carcinoma		
PT	comprises a nucleotide sequence.		
XX			
PS	Claim 1; SEQ ID NO 1709; Opp; English.		
XX			
CC	The present invention relates to novel DNA and protein sequences which		
CC	are associated with carcinomas. The sequences are useful for: (i) for		
CC	screening drug candidates; (ii) for screening of bioactive agent capable		
CC	of binding to Carcinoma Associated Protein (CAP); (iii) for screening of		
CC	a bioactive agent capable of modulating the activity of CAP; (iv) for		
CC	evaluating the effect of a candidate carcinoma drug; (v) for diagnosing		
CC	carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating		
CC	carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;		
CC	(x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for		
CC	determining Carcinoma Associated (CA) gene copy number. In addition, the		
CC	CA genes are useful as DNA vaccines and the CAP are useful as markers of		
CC	carcinoma including lymphoma. The present sequence is one such CA gene		
CC	sequence. Note: This patent is an equivalent to basic patent		
CC	US2002182586A1, for which no sequence data was published		
XX			
SQ	Sequence 5252 BP; 1507 A; 1101 C; 1177 G; 1467 T; 0 U; 0 Other;		
Query Match			
Best Local Similarity 75.2%; Score 1132.6; DB 11; Length 5252;			
Matches 1180; Conservative 3; Mismatches 28; Indels 3; Gaps 3;			
Qy	296	TGCCCCAACGGAACTTTCAGATTATATGCAATCGAAAGTGACAGGAACATGCCAACTCAA	355
Db	1	TGCCCAACGGAACTTTCAGATTATATGCAATCGAAAGTGACAGGAACATGCCAACTCAA	60
Qy	356	TCCCTCTTAAATGTACATGGAT-GGCCAAGAGTGAATGGCAGCTCTCTTTGCCAGTCCCGATG	414
Db	61	TCCCTCTTAAATGTACATGGATGGCCAGAGTGAATGGCAGCTCTCTTTGCCAGTCCCGATG	120
Qy	415	GAGATGGA-GATGCTTGTCAATGAAGGGCCNCTGTTGCAATTCGAGCTACACAA	473

121	Db	GAGATGAGGAGTGCCTTGTCTCAATGAAAGGGACCGCTGTGTTCATTTCCGAGCTACACAA	180
474	Qy	GAIAAAAAATGTCAATCCGAATCGAGGGGAATATCGCCCTTGGATTGCAATGTTCTCGACGCCA	533
181	Db	GAIAAAAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCAATGTTCTCGACGCCA	239
534	Qy	GACCTTCACACATTTCAAGAAGCTTAAATAAACATGTCTTAATGCACACACCGGCCCTACCCCT	593
240	Db	GACCTTCACACATTTCAAGAAGCTTAAATAAACATGTCTTAATGCACACACCGGCCCTACCCCT	299
594	Qy	CTGTGACCCAGCAGTTCCTTCGGGTGTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCA	653
300	Db	CTGTGAAACCCAGCAGTTCCTTCGGGTGTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCA	359
654	Qy	AGTCGGAACAGAACTCCCAAGGAAAAAGAAATGCAAGGAAAAATGAAATTTAGCTCTGAGGT	713
360	Db	AGTCGGAACAGAACTCCCAAGGAAAAAGAAATGCAAGGAAAAATGAAATTTAGCTCTGAGGT	419
714	Qy	ATGTGGCGAGACATTTTAGATGCGCTTTTGATGTGTGAGATCCACATGAGAAACACAAAGA	777
420	Db	ATGTGGCGAGACATTTTAGATGCGCTTTTGATGTGTGAGATCCACATGAGAAACACAAAGA	479
774	Qy	TTCTTTTCACTTACGGGTGTAACTGTGCGGAAGAAAGATTCAGAGAGCCTTGGTTTCTTAA	833
480	Db	TTCTTTTCACTTACGGGTGTAACTGTGCGGAAGAAAGATTCAGAGAGCCTTGGTTTCTTAA	539
834	Qy	AAATCACAATCGGACRCAATAATGGCAATCGGGGGCCAGAAAGCAAACTGCACGAAGGCTT	893
540	Db	AAATCACAATCGGAGACAATAATGGCAATCGGGGGCCAGAAAGCAAACTGCACGAAGGCTT	599
894	Qy	GGAGAGTAGTCCAGCAACGATCAACAGAGTCTCAGGTGCACGCGCCGAGAGCATCTC	953
600	Db	GGAGAGTAGTCCAGCAACGATCAACAGAGTCTCAGGTGCACGCGCCGAGAGCATCTC	659
954	Qy	CTCTCTTGCBAATCTGCATGTTTGTGCTTCTTATTTCCAAATAAGAAAGTCTTAAT	1013
660	Db	CTCTCTTGCBAATCTGCATGTTTGTGCTTCTTATTTCCAAATAAGAAAGTCTTAAT	719
1014	Qy	TGAGCAGCGCAAGGTGCACACCAAAAAAACTGCTTTCCGATACCAAGCAGCGCGACAGACA	1073
720	Db	TGAGCAGCGCAAGGTGCACACCAAAAAAACTGCTTTCCGATACCAAGCAGCGCGACAGACA	779
1074	Qy	CTCTCCACAAGGAGGAATGCGTCTCTCGAGGAGGACTTCTCTGAGTGTTCACATTGAG	1133
780	Db	CTCTCCACAAGGAGGAATGCGTCTCTCGAGGAGGACTTCTCTGAGTGTTCACATTGAG	839
1134	Qy	ACCAAAATCTCACCTGAAACGGGGAGAGCCTGTGAGATGCATCCCTCAGCTCGATCC	1193
840	Db	ACCAAAATCTCACCTGAAACGGGGAGAGCCTGTGAGATGCATCCCTCAGCTCGATCC	899
1194	Qy	GTTCCACCACCTTCAGGCTTTGGCAKCTGGCTTACCAAAGGAAGAGTGTGCCATTTGCCAAGA	1253
900	Db	GTTCCACCACCTTCAGGCTTTGGCAKCTGGCTTACCAAAGGAAGAGTGTGCCATTTGCCAAGA	959
1254	Qy	AGTGAAGGAATTTGGGCAAGAAGGGAGCACCGACAACGACGATTCGAGTTCGGAAGAAGGA	1313
960	Db	AGTGAAGGAATTCGGGGCAAGAAGGGAGCACCGACAACGACGATTCGGAAGAAGGA	1019
1314	Qy	GCTTTGGAGAAACAAATAAGAAACCTTTGTGAGGGCTCTCCGACAGAGAAAGAGAGTCCAA	1373
1020	Db	GCTTTGGAGAAACAAATAAGGGCAGTTGTGACGGCTCTCCGACAGAGAAAGAGAGTCCAA	1079
1374	Qy	ACACTCCACGGCGAAGCGCCCTCCGTGGAGCGGGATCCCAAGTTACCCAGTAGCAAGGA	1433
1080	Db	ACACTCCACGGCGAAGCGCCCTCCGTGGAGCGGGATCCCAAGTTACCCAGTAGCAAGGA	1139
1434	Qy	GAAGCCCACTCACTGCTCCGAGTGCGGCAAGACTTTTCAGAACCTTACCAACAGCTGTGCTT	1493
1140	Db	GAAGCCCACTCACTGCTCCGAGTGCGGCAAGACTTTTCAGAACCTTACCAACAGCTGTGCTT	1199
1494	Qy	GCACCTCAGGGTCC	1507
1200	Db	GCACCTCAGGGTCC	1213

RESULT 10
AAV09024
ID AAV09024 standard; DNA; 3183 BP.
XX AC
XX AAV09024;
XX
DT 21-JUL-1998 (first entry)
XX
XX Homo sapiens 20q13 amplicon ZABC-1 cDNA sequence.
XX
XX 20q13 amplicon; chromosome 20; tumour; detection; ZABC-1 gene;
KW chromosomal abnormalities; probe; gene therapy; antisense inhibition;
KW treatment; age-related macular degeneration; retinitis pigmentation;
KW Leber's congenital amaurosis; zinc finger amplified in breast cancer; ds.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 1..3183
FT CDS /*tag= a
FT /*product= "ZABC1 protein"
XX
XX WO9802539-A1.
XX
XX 22-JAN-1998.
XX
XX 15-JUL-1997; 97WO-US012343.
XX
XX 15-JUL-1996; 96US-00680395.
XX 16-OCT-1996; 96US-00731499.
XX 17-JAN-1997; 97US-00785532.
XX
XX (REGC) UNIV CALIFORNIA.
XX
XX Gray JW, Collins CC, Hwang S, Godfrey T, Kowbel D, Rommens J;
XX WPI; 1998-110587/10.
XX P-PSDB; AAW23975.
XX
XX New sequences from the 20q13 amplicon - used for detecting chromosomal
PT abnormalities, particularly tumours, and for developing products for
PT treating diseases.
XX
XX Claim 1; Page 67-69; 91pp; English.
XX
XX The sequence is that of the cDNA sequence encoding ZABC-1 (zinc finger
CC amplified in breast cancer). It maps to the core of the 20q13.2 amplicon
CC and is overexpressed in primary tumours and breast cancer cell lines
CC having 20q13.2 amplification. The sequence can be used as a probe for the
CC detection of chromosomal abnormalities at 20q13. It and other sequences
CC isolated from the 20q13 amplicon are consistently amplified in primary
CC tumours. These sequences are useful as probes or as probe targets for
CC monitoring the relative copy number of corresponding sequences from a
CC biological sample such as tumour cells. The sequences can also be used in
CC therapeutic applications for modulating the expression of the endogenous
CC gene or the activity of the gene product. Examples of therapeutic
CC approaches include antisense inhibition of gene expression, gene therapy,
CC and monoclonal antibodies that specifically bind the gene products. The
CC products can also be used in the treatment of other diseases, e.g. age-
CC related macular degeneration, Leber's congenital amaurosis and retinitis
CC pigmentata
XX
SQ Sequence 3183 BP; 925 A; 796 C; 758 G; 698 T; 0 U; 6 Other;

Query Match 72.1%; Score 1087.2; DB 2; Length 3183;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1138; Conservative 9; Mismatches 38; Indels 4; Gaps 3

Qy 320 ATGCAATCGAAAGTGACAGGAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGC 379
Db 1 ATGCAATCGAAAGTGACAGGAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGG 60

QY	1459	GGCAAGCTTTTCAGAACCTACCAACGAGCTGGTCTTGCATCTCAGGGTCC	1507
Db	1138	GGCAAGCTTTTCAGAACCTACCAACGAGCTGGTCTTGCATCTCAGGGTCC	1186
RESULT 11			
ACN44984			
ID	ACN44984	standard; DNA; 26345 BP.	
XX	XX	ACN44984;	
AC	ACN44984;		
DT	18-NOV-2004	(first entry)	
XX	XX	Mouse genomic sequence MCG6549.	
DE	XX	Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.	
KW	XX	Mus musculus.	
OS	XX	WO2003073826-A2.	
PN	XX	12-SEP-2003.	
XX	XX	28-FEB-2003; 2003WO-US006235.	
PP	XX	01-MAR-2002; 2002US-00087192.	
PR	XX	(SAGR-) SAGRES DISCOVERY.	
PA	XX	Morris DW;	
XX	XX	WPI; 2003-328604/31.	
DR	XX	Recombinant nucleic acid useful for diagnosis and treatment of carcinoma	
XX	XX	comprises a nucleotide sequence.	
PT	XX	Claim 1; SEQ ID NO 1705; Opp; English.	
PS	XX	The present invention relates to novel DNA and protein sequences which	
CC	CC	are associated with carcinomas. The sequences are useful for: (i) for	
CC	CC	screening drug candidates; (ii) for screening of bioactive agent capable	
CC	CC	of binding to Carcino Associated Protein (CAP); (iii) for screening of	
CC	CC	a bioactive agent capable of modulating the activity of CAP; (iv) for	
CC	CC	evaluating the effect of a candidate carcinoma drug; (v) for diagnosing	
CC	CC	carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating	
CC	CC	carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;	
CC	CC	(x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for	
CC	CC	determining Carcino Associated (CA) gene copy number. In addition, the	
CC	CC	CA genes are useful as DNA vaccines and the CAP are useful as markers of	
CC	CC	carcinoma including lymphoma. The present sequence is one such CA coding	
CC	CC	sequence. Note: This patent is an equivalent to basic patent	
CC	CC	US2002182586A1, for which no sequence data was published	
XX	SQ	Sequence 26345 BP; 6286 A; 5855 C; 6718 G; 7486 T; 0 U; 0 Other;	
Query Match			
Best Local Similarity 38.1%; Score 573.6; DB 11; Length 26345;			
Matches 1028; Conservative 3; Mismatches 448; Indels 88; Gaps 10;			
QY	2	CAGGTTGCTGGAGTTGACTTCTTGTCTCAATTTGAACACTTTCATTCATGAGACAAAGAGC	61
Db	9674	CACATTCGCGGATCAACGCTCTGTCTCACTGAGCCACTCACCAGTGGAGATGGACAGAG	9733
QY	62	ACTAATGCTTGTGCTGATTTCATATTTGAATCGAGCATTTGGNACCCCTGTATGCTTGT	121
Db	9734	CTGTGCTTGTGCTGAGCACTGACTCAGCCCTGGGGCAGCTTGGATGCTTGT	9793
QY	122	TTGTGGAAGAACCACTGACACCATCACTGAGCTTCTCTAAAGTTTCGAAGAAGTTAGAGG	181
Db	9794	TTGTGCAAGAGCCA---TCTCCATCACTGAGCTCTCCAGCTCC-----GG	9838
QY	182	ACTATACACTTTCTTTTGAACCTTTTATAATAATATTTGCTCTGTTTGGAAACCCAGGA	241

Db	9839	AATATACCTGCTCCCTCTGACCTT-----AGAGGTGCTCTGTATTTTGAACCCAGGG	9889
QY	242	CTGTTTAGAGGGTGAG---TGACAGGTCTTACAGTGGCCCTTAATCAAACTCCAGAAATTGC	298
Db	9890	CTCTTAGAGGTTCTGAGTGGTGATATCTCACAGGCGGCTTATTTCTAATGTCAGAGATCC	9949
QY	299	CCAAACGGAACCTTTGAGATTATATGCAATCGA-----AAGTCACAGGAACAACTGCCAAT	352
Db	9950	TCAGCAGATAGTGGATTATCCCATCTCTGTACAGACCCCTGGAAGTGGTGGCAGCATGCCGAC	10009
QY	353	CAATCCCTCTTAATGTACATGGA--TGGCCAAAGAGTATGGCAGCTCTCTTCCAGTCCG	411
Db	10010	CAGTCCCTCTCTGTATCATGACGCGGCGGAAGTCTCTCAGAGCTCTCTAGCTCCCAAG	10069
QY	412	ATGGAGATGGAGATGCCCTTGTCAATCAAAAGGCCCNCCTGTTGTCTCAATTCGAGCTACACA	471
Db	10070	ATGGAGTGGATGATGCTGTGCCATAAAAGGCGGTGGCAGTCCCTTCCGAGCTGT	10129
QY	472	AAGAAAAAATGTCAATTCGGAATCGAGGGAATATGCCCTTGGATTCGATGTCGAGC	531
Db	10130	CAGGAGAAAGCATGGCCGTGGCAGAGGCCACATGCCCTCGATTCGATGCTTCTGAGC	10189
QY	532	CAGACCTTCACACATTCAGAGACCTTAAATAACATGCTTAAATGCACACCGGCTTACC	591
Db	10190	CAGGCTCTCTCTCAGGCGGAGATCTCAGTCAGCAGCTGCTCTGAGCACCGGCCACC	10249
QY	592	CTCTGTGAACACAGCAGTCTTTCGGGTGGAAGCAGAGTATCTCAGTCCGCTTGATAAAAAGT	651
Db	10250	CTCTCGAGCCAGCTGCTCTGCTGTGGAGCGGAGTACCTAAGTCCCTTGTATTAAGCT	10309
QY	652	CAAGTGCAGAACAGAACCTCCCAAGAAAAAGAAATTCGAAGGA---AAATGAATTTAGCTGT	708
Db	10310	CTGGAGCCACACAGAGCCAGCATTTGGAGAAGAGTGGCGAAGACCCCGAGGAGTTGAGCTGT	10369
QY	709	GAGGTATGTGGCAGACACATTTAGACTCGCTTTTGTGATTTGAGATCCACATGAGAACACAC	768
Db	10370	GATGTGTGGGCAGACATTTCCAGTGGCTTTGTGATTTGAGAGCCACATGAAGAAGCAT	10429
QY	769	AAAGATTTCTTTCACATTACGGGTGTAACATGTGCGGAAGAAGATTCAAGAGAGCTTGGTTT	828
Db	10430	AAGGACTCTTTCACGTATGGGTGAGCATGTGCGGAGGAGGATTCAGAGGCGGTGTTTC	10489
QY	829	CTTAAAAATCAATCGGACACATTAATGCAAAATCGGGGGCCAGAAAGCAAACTGCAGCAA	888
Db	10490	CTGAAGAACCACATCGGACACACAATGCAAGTCTGGCACCAGGAGCAAGCTTCAGCAA	10549
QY	889	GCCTTGGAGATGATCTCAGCAACGATCAACAGGTCTCTCAGTGCACCGCGCCGAGAGC	948
Db	10550	GGCATGGAG---AGTCCAGTCAACATCAATGAAGTGGTCCAGCCGACGCCCTGGGAGC	10606
QY	949	ATCTCTCTCTCTTGCAAAATCTGCATGGTTTGTGGCTTCTTATTTCCAAATAAAGAAAGT	1008
Db	10607	ATCTCAGGCCCTACAGATCTGCATGGTCTGCGGCTTCTCTTCCCAANTAAAGCAGAGC	10666
QY	1009	CTAATTGAGCACCAGGTGCACACCAAAAAAACTGCTTTTCGTTCCAGCAGCGCGCAG	1068
Db	10667	CTCATTGAGCAGCAAGGTTCAAGCAAAAGAACTGTCTCCAGTGCAGCAACGTTGCC	10726
QY	1069	ACAGACTCTCCACAGGAGGAATGCCCTCTCGAGGAGGAGCTTCTGCAAGTTGTTCAAC	1128
Db	10727	CCTGATGATCCACCGAGAGAAACCCACGCTCCCGAGGGAAGTGTGTCAGTTTTTGAAC	10786
QY	1129	TTGAGACCAAAATCTCACCTGAAACGGGAAGAAGCTGTGATGATGATCTCAGCTC	1188
Db	10787	TTGAGACCAAGATCAATGTCAGGTAGTACAGTGAAGCCCATGACCTGCATACCTAGCTT	10846
QY	1189	GATCGCTTCCACACCTTCCAGGCTTGGCAKCTGGCTTACCAAGAAAGTGTGCCATTTGC	1248
Db	10847	GACCCGTTCCACACCTACAGGATGGCAGTGGCTTACCAAGAAAGTGTGCCGCTTGC	10906
QY	1249	C---AAGAAGTGAAGAAATTTGGGGCAAGAGGGAGCAGCAACGACGATTCGAGTTCC	1305
Db	10907	CAGGAAGAGGTGAAAGAGTCAAGCCCAAGAGGAAGCAGACAATGACGACTCATGCTCA	10966

Db	1018	GGAAAGTCCAAACAAGTAAAGCAGTTGTCCAGGTCTCTCCCAAGCAAGAGAGCCT	1077
QY	1372	AAACACTCCACGGCAGCGCCTCCGTGGACGGGATCCCAAGTTACCCAGTAGCAAG	1431
Db	1078	AGACATGCTAATAGTAGTGAAGTGCCCTTCTGGGGATAGTGACCCCAAGTTGTCCAGTAGCAAG	1137
QY	1432	GAGAAGCCCACTCACTGCTCCGAGTGGGCAAGCTTTCAGAACCTTACCACCAAGCTGGTC	1491
Db	1138	GAGAAGCCCACTCACTGCTTCTGAGTGCAGCAAGCCTTCAGGACATACCACCAAGCTCGTC	1197
QY	1492	TTGCACTCCAGGGTCC	1507
Db	1198	CTGCACTCGAGGTGC	1213
RESULT 13			
AAK81095			
ID	AAK81095 standard; DNA; 566 BP.		
AC	AAK81095;		
XX			
DT	07-NOV-2001 (first entry)		
XX			
DE	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35907.		
XX			
KW	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;		
KW	cytostatic; gene therapy; vaccine; metastasis; ds.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200157182-A2.		
XX			
PD			
XX			
PF	09-AUG-2001.		
XX			
17-JAN-2001;	2001WO-US001354.		
XX			
31-JAN-2000;	2000US-0179065P.		
PR			
04-FEB-2000;	2000US-0180628P.		
PR			
24-FEB-2000;	2000US-0184664P.		
PR			
02-MAR-2000;	2000US-0186350P.		
PR			
16-MAR-2000;	2000US-0189874P.		
PR			
17-MAR-2000;	2000US-0190076P.		
PR			
18-APR-2000;	2000US-0198123P.		
PR			
19-MAY-2000;	2000US-0205515P.		
PR			
07-JUN-2000;	2000US-0209467P.		
PR			
28-JUN-2000;	2000US-0214886P.		
PR			
30-JUN-2000;	2000US-0215135P.		
PR			
07-JUL-2000;	2000US-0216647P.		
PR			
07-JUL-2000;	2000US-0216880P.		
PR			
11-JUL-2000;	2000US-0217487P.		
PR			
11-JUL-2000;	2000US-0217496P.		
PR			
14-JUL-2000;	2000US-0218290P.		
PR			
26-JUL-2000;	2000US-0220963P.		
PR			
26-JUL-2000;	2000US-0220964P.		
PR			
14-AUG-2000;	2000US-0224518P.		
PR			
14-AUG-2000;	2000US-0224519P.		
PR			
14-AUG-2000;	2000US-0225213P.		
PR			
14-AUG-2000;	2000US-0225214P.		
PR			
14-AUG-2000;	2000US-0225266P.		
PR			
14-AUG-2000;	2000US-0225267P.		
PR			
14-AUG-2000;	2000US-0225268P.		
PR			
14-AUG-2000;	2000US-0225270P.		
PR			
14-AUG-2000;	2000US-0225447P.		
PR			
14-AUG-2000;	2000US-0225757P.		
PR			
14-AUG-2000;	2000US-0225758P.		
PR			
18-AUG-2000;	2000US-0226279P.		
PR			
22-AUG-2000;	2000US-0226681P.		
PR			
22-AUG-2000;	2000US-0226688P.		
PR			
22-AUG-2000;	2000US-0227182P.		
PR			
23-AUG-2000;	2000US-0227009P.		
PR			
30-AUG-2000;	2000US-0228924P.		
PR			
01-SEP-2000;	2000US-0229287P.		
PR			
01-SEP-2000;	2000US-0229343P.		
PR			
01-SEP-2000;	2000US-0229344P.		
PR			
01-SEP-2000;	2000US-0229345P.		
PR			
05-SEP-2000;	2000US-0229509P.		
PR			
05-SEP-2000;	2000US-0229513P.		
PR			
06-SEP-2000;	2000US-0230437P.		
PR			
06-SEP-2000;	2000US-0230438P.		
PR			
08-SEP-2000;	2000US-0231242P.		
PR			
08-SEP-2000;	2000US-0231243P.		
PR			
08-SEP-2000;	2000US-0231244P.		
PR			
08-SEP-2000;	2000US-0231413P.		
PR			
08-SEP-2000;	2000US-0231414P.		
PR			
08-SEP-2000;	2000US-0232080P.		
PR			
08-SEP-2000;	2000US-0232081P.		
PR			
12-SEP-2000;	2000US-0231968P.		
PR			
14-SEP-2000;	2000US-0232397P.		
PR			
14-SEP-2000;	2000US-0232398P.		
PR			
14-SEP-2000;	2000US-0232399P.		
PR			
14-SEP-2000;	2000US-0232400P.		
PR			
14-SEP-2000;	2000US-0232401P.		
PR			
14-SEP-2000;	2000US-0233063P.		
PR			
14-SEP-2000;	2000US-0233064P.		
PR			
14-SEP-2000;	2000US-0233065P.		
PR			
21-SEP-2000;	2000US-0234223P.		
PR			
21-SEP-2000;	2000US-0234274P.		
PR			
25-SEP-2000;	2000US-0234997P.		
PR			
25-SEP-2000;	2000US-0234998P.		
PR			
26-SEP-2000;	2000US-0235484P.		
PR			
27-SEP-2000;	2000US-0235834P.		
PR			
27-SEP-2000;	2000US-0235836P.		
PR			
29-SEP-2000;	2000US-0236327P.		
PR			
29-SEP-2000;	2000US-0236367P.		
PR			
29-SEP-2000;	2000US-0236368P.		
PR			
29-SEP-2000;	2000US-0236369P.		
PR			
29-SEP-2000;	2000US-0236370P.		
PR			
02-OCT-2000;	2000US-0236802P.		
PR			
02-OCT-2000;	2000US-0237037P.		
PR			
02-OCT-2000;	2000US-0237038P.		
PR			
02-OCT-2000;	2000US-0237039P.		
PR			
13-OCT-2000;	2000US-0237040P.		
PR			
13-OCT-2000;	2000US-0239935P.		
PR			
13-OCT-2000;	2000US-0239937P.		
PR			
20-OCT-2000;	2000US-0240960P.		
PR			
20-OCT-2000;	2000US-0241221P.		
PR			
20-OCT-2000;	2000US-0241785P.		
PR			
20-OCT-2000;	2000US-0241786P.		
PR			
20-OCT-2000;	2000US-0241787P.		
PR			
20-OCT-2000;	2000US-0241808P.		
PR			
20-OCT-2000;	2000US-0241809P.		
PR			
20-OCT-2000;	2000US-0241826P.		
PR			
01-NOV-2000;	2000US-0244617P.		
PR			
08-NOV-2000;	2000US-0246474P.		
PR			
08-NOV-2000;	2000US-0246475P.		
PR			
08-NOV-2000;	2000US-0246476P.		
PR			
08-NOV-2000;	2000US-0246478P.		
PR			
08-NOV-2000;	2000US-0246523P.		
PR			
08-NOV-2000;	2000US-0246524P.		
PR			
08-NOV-2000;	2000US-0246525P.		
PR			
08-NOV-2000;	2000US-0246528P.		
PR			
08-NOV-2000;	2000US-0246527P.		
PR			
08-NOV-2000;	2000US-0246529P.		
PR			
08-NOV-2000;	2000US-0246532P.		
PR			
08-NOV-2000;	2000US-0246533P.		
PR			
08-NOV-2000;	2000US-0246609P.		
PR			
08-NOV-2000;	2000US-0246610P.		
PR			
08-NOV-2000;	2000US-0246611P.		
PR			
08-NOV-2000;	2000US-0246613P.		
PR			
17-NOV-2000;	2000US-0249207P.		
PR			
17-NOV-2000;	2000US-0249208P.		
PR			
17-NOV-2000;	2000US-0249209P.		
PR			
17-NOV-2000;	2000US-0249210P.		
PR			
17-NOV-2000;	2000US-0249211P.		
PR			
17-NOV-2000;	2000US-0249212P.		
PR			

Query Match 21.5%; Score 324.2; DB 4; Length 14906;
Best Local Similarity 97.9%; Pred. No. 2e-93; Mismatches 2; Indels 0; Gaps 0;
Matches 326; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 1175 GCATCCCTCAGTCGATCGTTCACCACTTCCAGGCTTGGCAKCTGGCTACCAAGGAA 1234
DB 1 GCATCCCTCAGTCGATCGTTCACCACTTCCAGGCTTGGCAGCTGGCTACCAAGGAA 60
QY 1235 WAGTTGCCATTGCGCAAGTGAAGGAATTTGGGCGAAGAGGAGCACCACACGACG 1294
DB 61 AAGTTGCCATTGCGCAAGTGAAGGAATCGGGCGCAAGAGGAGCACCACACGACG 120
QY 1295 ATTGAGTTCCGAGAAAGAGCTTGGAGAAACAAATAAGAACCATTTGTCAGGCTCTCGC 1354
DB 121 ATTGAGTTCCGAGAAAGAGCTTGGAGAAACAAATAAGGAGCATTTGTCAGGCTCTCGC 180
QY 1355 AAGAGAAAGAGAAAGTGCACCACTCCACGCGGAAGCCCTCCGTGGACGCGGATCCCA 1414
DB 181 AAGAGAAAGAGAAAGTGCACCACTCCACGCGGAAGCCCTCCGTGGACGCGGATCCCA 240
QY 1415 AGTTACCCAGTAGCAGAGAGAGCCCACTCACTGCTCCGAGTGGCGGCAAGCTTTCAGAA 1474
DB 241 AGTTACCCAGTAGCAGAGAGAGCCCACTCACTGCTCCGAGTGGCGGCAAGCTTTCAGAA 300
QY 1475 CCTACACAGCTGCTCTTGCACCTCCAGGGTCC 1507
DB 301 CCTACACAGCTGCTCTTGCACCTCCAGGGTCC 333

RESULT 15
ACH76516/c
ID: ACH76516 standard; DNA; 530 BP.
XX ACH76516;
AC ACH76516;
XX 29-JUL-2004 (first entry)
XX Human genome derived single exon probe #9711.
DE Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX Homo sapiens.
XX US2003194704-A1.
XX 16-OCT-2003.
XX 03-APR-2002; 2002US-00029386.
XX 03-APR-2002; 2002US-00029386.
XX (PENN/) PENN S G.
XX (RANK/) RANK D R.
XX (HANZ/) HANZEL D K.
XX Penn SG, Rank DR, Hanzel DK;
XX WPI; 2004-119264/12.
XX New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX Claim 15; SEQ ID NO 9711; 80pp; English.
XX The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 688 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule

expressed in human cells or tissues. Also included are a spatially-
addressable set of single exon nucleic acid probes for measuring human
gene expression (comprising a plurality of single exon nucleic acid
probes cited above, where each of the plurality of probes is separately
and addressably isolatable or amplifiable from the plurality), a single
exon microarray for measuring human gene expression, a method of
measuring human gene expression, a vector comprising the single exon
probe cited above, an ORF-encoded peptide comprising at least 8
contiguous amino acids of any of the above-mentioned amino acid
sequences (optionally with conservative amino acid substitutions), an
isolated antibody that binds specifically to a peptide cited above,
methods of selling and/or licensing single exon probes or microarrays to
a customer desiring to measure gene expression, a method of providing
human gene expression data by subsequence, and a computer-readable
storage medium which contains a database having a plurality of records
(each record including data on the expression of a single exon probe
cited above). The probe, methods and apparatus are useful in gene
expression analysis. The probes may be used as tools for surveying
tissues to detect the presence of expressed messages that contain their
specific exon, or in constructing genome-derived single exon microarrays.
In addition, the probes are used in identifying and characterising
alternative splicing events, in detecting and characterising gross
alterations in the genomic locus that includes their exon, in assessing
smaller genomic alterations, in priming the synthesis of nucleic acids,
or in expressing the ORF-encoded peptide. The present sequence is a human
single exon probe of the invention. Note: The sequence data for this
patent did not form part of the printed specification, but was obtained
in electronic format directly from USPTO at
seqdata.uspto.gov/sequence.html?docID=20030194704
XX
SQ Sequence 530 BP; 97 A; 160 C; 156 G; 117 T; 0 U; 0 Other;

Query Match 14.3%; Score 215.6; DB 12; Length 530;
Best Local Similarity 98.2%; Pred. No. 6.1e-59;
Matches 218; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1286 ACAACGACGATTTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGAACCATTTGTGCGAG 1345
DB 530 ACAACGACGATTTCGAGTTCCGAGAAGGAGCTTGGAGAAACAAATAAGGGCAGTTGTGCGAG 471
QY 1346 GCCTCTCGCAAGAGAGAAAGAGAGTGCAAACTCCACGGCGAAGCGCCCTCCGTGGACG 1405
DB 470 GCCTCTCGCAAGAGAGAAAGAGAGTGCAAACTCCACGGCGAAGCGCCCTCCGTGGACG 411
QY 1406 CGGATCCCAAGTTACCCAGTAGCAGAGAGAGCCCACTCACTGCTCCGAGTGGCGCAAG 1465
DB 410 CGGATCCCAAGTTACCCAGTAGCAGAGAGAGCCCACTCACTGCTCCGAGTGGCGCAAG 351
QY 1466 CTTTCAGAACCTTACCACGAGCTGGTCTTGGCACTCCAGGGTCC 1507
DB 350 CTTTCAGAACCTTACCACGAGCTGGTCTTGGCACTCCAGGGTCC 309

Search completed: June 28, 2005, 19:33:39
Job time : 537.33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 18:22:52 ; Search time 171.067 Seconds
(without alignments)
14414.621 Million cell updates/sec

Title: US-08-731-499-3
Perfect score: 1507
Sequence: 1 GCAGGTCTGCTGGATTGACT.....GGCTTGCACCTCCAGGGTCC 1507

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B COMB.seq: *
5: /cgn2_6/prodata/1/ina/PTUS COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1504.8	99.9	1507	2	US-08-680-395-3
2	1504.8	99.9	1507	4	US-08-892-695-3
3	1387.2	92.1	10365	4	US-08-892-695-9
4	1387.2	92.1	20022	4	US-09-949-016-12604
5	1387.2	92.1	20023	4	US-09-949-016-16004
6	1337.2	88.7	5632	3	US-09-560-594-3
7	1337.2	88.7	5632	4	US-09-949-016-862
8	1337.2	88.7	5632	4	US-09-949-016-4262
9	1108.6	73.6	3186	4	US-08-892-695-10
10	550.8	36.5	601	4	US-09-949-016-34699
11	550.8	36.5	601	4	US-09-949-016-151879
12	454.8	30.2	601	4	US-09-949-016-34700
13	454.8	30.2	601	4	US-09-949-016-151880
14	197.2	13.1	298	4	US-09-513-999C-11828
15	68.8	4.6	2765	4	US-09-620-312D-61
16	68.8	4.6	2799	4	US-09-949-016-4961
17	68.8	4.6	3039	4	US-09-949-016-1038
18	68.8	4.6	12323	4	US-09-949-016-16703
19	50.2	3.3	7218	1	US-08-232-463-14
20	48.6	3.2	601	4	US-09-949-016-27493
21	48.6	3.2	601	4	US-09-949-016-92521
22	48.6	3.2	1460	4	US-09-774-528-413
23	48.6	3.2	1820	4	US-09-949-016-2645
24	48.6	3.2	1839	4	US-09-949-016-493
25	48.6	3.2	23174	4	US-09-949-016-14387
26	48.6	3.2	23187	4	US-09-949-016-12235
27	48	3.2	2978	4	US-09-949-016-3823

28	48	3.2	15632	4	US-09-949-016-15565	Sequence 15565, A
29	47.2	3.1	1563	4	US-09-016-434-1386	Sequence 1386, Ap
30	47	3.1	2752	4	US-09-949-016-3563	Sequence 3563, Ap
31	47	3.1	18824	4	US-09-949-016-15305	Sequence 15305, A
32	46.8	3.1	831	4	US-09-016-434-725	Sequence 725, App
33	46.4	3.1	656	4	US-09-016-434-1406	Sequence 1406, Ap
34	46.2	3.1	857	4	US-09-016-434-892	Sequence 892, App
35	44.2	2.9	605	4	US-09-621-976-2809	Sequence 2809, Ap
36	44	2.9	293	4	US-09-016-434-23	Sequence 23, Appl
37	43.4	2.9	2402	4	US-09-949-016-1710	Sequence 1710, Ap
38	43.4	2.9	2407	4	US-09-949-016-489	Sequence 489, App
39	43.4	2.9	30867	4	US-09-949-016-12231	Sequence 12231, A
40	43.4	2.9	30868	4	US-09-949-016-13452	Sequence 13452, A
41	43.2	2.9	2555	4	US-09-620-312D-1050	Sequence 1050, Ap
42	43	2.9	241	4	US-09-016-434-746	Sequence 746, App
43	43	2.9	2920	4	US-09-620-312D-1084	Sequence 1084, Ap
44	41.4	2.7	1988	4	US-09-620-312D-1080	Sequence 1080, Ap
45	40.6	2.7	1189	4	US-09-016-434-1400	Sequence 1400, Ap

ALIGNMENTS

RESULT 1
US-08-680-395-3
; Sequence 3, Application US/08680395
; Patent No. 5892010
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W.
; APPLICANT: Collins, Colin
; APPLICANT: Hwang, Soo-in
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowbel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: Genes from the 20q13 Amplicon and Their
; TITLE OF INVENTION: Uses
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/680,395
; FILING DATE: 15-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-06890005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1507 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: 1..1507
; LOCATION: /note="cDNA clone cc49 of 6-7kb
; OTHER INFORMATION: transcript with homology to C2H2 zinc
; OTHER INFORMATION: finger genes"

US-08-680-395-3

Query Match 99.9%; Score 1504.8; DB 2; Length 1507;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1507; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGTTTCTGGGATGACTTCTTCTCAATGAACACTCAATTCGAGGACCAAGAG 60
 Db 1 GCAGTTTCTGGGATGACTTCTTCTCAATGAACACTCAATTCGAGGACCAAGAG 60

QY 61 CACTAATGCTTTGCTGCTCAATTCATATTTGAATCGAGGCAATGGGAACCTGTATGCCCTTG 120
 Db 61 CACTAATGCTTTGCTGCTCAATTCATATTTGAATCGAGGCAATGGGAACCTGTATGCCCTTG 120

QY 121 TTTGTGGAAGAACCAAGTGACACCACTACTGAGCTTCTTAAAGTTTCGAAGAAGTTAGAG 180
 Db 121 TTTGTGGAAGAACCAAGTGACACCACTACTGAGCTTCTTAAAGTTTCGAAGAAGTTAGAG 180

QY 181 GACTATACACTTTCTTTTGAACCTTTTATATTAATATTTGCTCTGCTTTTGGAAACCCAGG 240
 Db 181 GACTATACACTTTCTTTTGAACCTTTTATATTAATATTTGCTCTGCTTTTGGAAACCCAGG 240

QY 241 ACTGTTAGAGGTGAGTGACAGGCTCTTACAGTGGCTTTAATCCAACTCCAGAAATTTGCC 300
 Db 241 ACTGTTAGAGGTGAGTGACAGGCTCTTACAGTGGCTTTAATCCAACTCCAGAAATTTGCC 300

QY 301 AACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAACATGCCAACTCAATCCCT 360
 Db 301 AACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAACATGCCAACTCAATCCCT 360

QY 361 CTTAATGTACATGATGATGCAAGAGTGATGCGAGCTCTCTTGGCCAGTCCGATGAGATG 420
 Db 361 CTTAATGTACATGATGATGCAAGAGTGATGCGAGCTCTCTTGGCCAGTCCGATGAGATG 420

QY 421 GAGATGCTTTGCAATGAAAGGGCCNCTGTGTGCAATTCGAGCTACACAAAGAAAAA 480
 Db 421 GAGATGCTTTGCAATGAAAGGGCCNCTGTGTGCAATTCGAGCTACACAAAGAAAAA 480

QY 481 ATGTCAATCCGAATCGAGGGGAATATGCCCTTTGGATTGATTTCTGAGGCGAGACCTTC 540
 Db 481 ATGTCAATCCGAATCGAGGGGAATATGCCCTTTGGATTGATTTCTGAGGCGAGACCTTC 540

QY 541 ACACATTCAGAGACCTTAATAACATGCTTAATGCAACACCGCCCTACCTCTGTGAA 600
 Db 541 ACACATTCAGAGACCTTAATAACATGCTTAATGCAACACCGCCCTACCTCTGTGAA 600

QY 601 CCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGTGGGA 660
 Db 601 CCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGTGGGA 660

QY 661 ACAGAACCTCCCAAGGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATGTGGG 720
 Db 661 ACAGAACCTCCCAAGGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATGTGGG 720

QY 721 CAGACATTTAGAGTGCCTTTTGAATGTTGAGATCCACATGAGAACACAAAGATTTCTTC 780
 Db 721 CAGACATTTAGAGTGCCTTTTGAATGTTGAGATCCACATGAGAACACAAAGATTTCTTC 780

QY 781 ACTTACGGGTGTAAATGTGCGGAAGAAAGATTCAGAGGAGCTTTGTTTCTTAAAAATCAC 840
 Db 781 ACTTACGGGTGTAAATGTGCGGAAGAAAGATTCAGAGGAGCTTTGTTTCTTAAAAATCAC 840

QY 841 ATGGGACRCATAATGCGAAATCGGGGCGAGAACCAACTGACAGCAAGGCTTGAGAGT 900
 Db 841 ATGGGACRCATAATGCGAAATCGGGGCGAGAACCAACTGACAGCAAGGCTTGAGAGT 900

QY 901 AGTCCAGCAACGATCAACGAGTCTGTCAGGTGACGGCCGAGAGCATCTCTCTCT 960
 Db 901 AGTCCAGCAACGATCAACGAGTCTGTCAGGTGACGGCCGAGAGCATCTCTCTCT 960

QY 961 TGCAAAATCTGCATGTTGTGGCTTCTCTATTTCCAAATTAAGAAAGTCTAATTGAGCAC 1020
 Db 961 TGCAAAATCTGCATGTTGTGGCTTCTCTATTTCCAAATTAAGAAAGTCTAATTGAGCAC 1020

RESULT 2

US-08-892-695-3
 ; Sequence 3, Application US/08892695A
 ; Patent No. 6808878
 ; GENERAL INFORMATION:
 ; APPLICANT: Gray, Joe W
 ; APPLICANT: Collins, Collin
 ; APPLICANT: Hwang, Soo In
 ; APPLICANT: Godfrey, Tony
 ; APPLICANT: Kowel, David
 ; APPLICANT: Rommens, Johanna
 ; TITLE OF INVENTION: GENES FROM THE 20Q13 AMPLICON AND THEIR USES
 ; FILE REFERENCE: 2500.124US3
 ; CURRENT APPLICATION NUMBER: US/08/892,695A
 ; CURRENT FILING DATE: 1997-07-15
 ; EARLIER APPLICATION NUMBER: 08/785,532
 ; EARLIER FILING DATE: 1997-01-17
 ; EARLIER APPLICATION NUMBER: 08/731,499
 ; EARLIER FILING DATE: 1996-10-16
 ; EARLIER APPLICATION NUMBER: 08/680,395
 ; EARLIER FILING DATE: 1996-07-15
 ; NUMBER OF SEQ ID NOS: 59
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 1507
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:cc49
 ; FEATURE:
 ; NAME/KEY: modified_base
 ; LOCATION: (447)
 ; OTHER INFORMATION: N is A, G, T, or U
 US-08-892-695-3

QY 1021 CGCAGGTGCAACCAAAAAAATGCTTTTCGGTACCAAGCGCCGACAGACTCTCCA 1080
 Db 1021 CGCAGGTGCAACCAAAAAAATGCTTTTCGGTACCAAGCGCCGACAGACTCTCCA 1080

QY 1081 CAAGGAGGAATGCCCTCTCGAGGGAGGACTTCCCTGCAAGTTGTTCAACTTGAGACCAAAA 1140
 Db 1081 CAAGGAGGAATGCCCTCTCGAGGGAGGACTTCCCTGCAAGTTGTTCAACTTGAGACCAAAA 1140

QY 1141 TCTCACTCCGTAAGCGGGGAAGAGCCTGTGAGATGCATCCCTCAGCTCGATCCGTTCAAC 1200
 Db 1141 TCTCACTCCGTAAGCGGGGAAGAGCCTGTGAGATGCATCCCTCAGCTCGATCCGTTCAAC 1200

QY 1201 ACCTTCAGGCTTGGCACTGCTACCAAGAAAGTTGCCATTTGCCAAGAAAGTGAAG 1260
 Db 1201 ACCTTCAGGCTTGGCACTGCTACCAAGAAAGTTGCCATTTGCCAAGAAAGTGAAG 1260

QY 1261 GAAATTTGGGGCAAGAGGAGGACCGACAAACGACGATTCGAGTTCCGAGAGAGCTTGA 1320
 Db 1261 GAAATTTGGGGCAAGAGGAGGACCGACAAACGACGATTCGAGTTCCGAGAGAGCTTGA 1320

QY 1321 GAAACAAATAAGAACCAATTTGTGAGGCTCTCGCAAGAGAAAGAGTGCACCACTCC 1380
 Db 1321 GAAACAAATAAGAACCAATTTGTGAGGCTCTCGCAAGAGAAAGAGTGCACCACTCC 1380

QY 1381 CACGGCGAAGGCCCTCGTGGACGCGGATCCCAAGTTACCCAGTAGCAAGAGAGGCC 1440
 Db 1381 CACGGCGAAGGCCCTCGTGGACGCGGATCCCAAGTTACCCAGTAGCAAGAGAGGCC 1440

QY 1441 ACTCACTCTCGAGTGGGCAAGCTTTTCAGAACCTTACCAAGCTGTGTGCACTCC 1500
 Db 1441 ACTCACTCTCGAGTGGGCAAGCTTTTCAGAACCTTACCAAGCTGTGTGCACTCC 1500

QY 1501 AGGGTCC 1507
 Db 1501 AGGGTCC 1507

Query Match		99.9%;	Score 1504.8;	DB 4;	Length 1507;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 1507;		Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
QY	1	GCAGGTTCGCGGATGACTCTCTGCTCAATTCGAACACTCATTCATTCGAGACAAAGAG	60		
DB	1	GCAGGTTCGCGGATGACTCTCTGCTCAATTCGAACACTCATTCATTCGAGACAAAGAG	60		
QY	61	CACATAATGCTTGTGCTGATTCATATTCGAATTCGAGGCAATTCGGAACCCCTGTATGCTCTTG	120		
DB	61	CACATAATGCTTGTGCTGATTCATATTCGAATTCGAGGCAATTCGGAACCCCTGTATGCTCTTG	120		
QY	121	TTTGTGGAAGAACCAAGTACACCTATCTAGCTGCTTCTTAAAGTTTCGAAGTTAGAG	180		
DB	121	TTTGTGGAAGAACCAAGTACACCTATCTAGCTGCTTCTTAAAGTTTCGAAGTTAGAG	180		
QY	181	GACTATACACTTCTTTTGAACCTTTTATATAATAATATTTGCTCTGCTTTCGAAACCCAGG	240		
DB	181	GACTATACACTTCTTTTGAACCTTTTATATAATAATATTTGCTCTGCTTTCGAAACCCAGG	240		
QY	241	ACTGTTAGAGGCTGAGTGACAGGCTCTTACAGTGGCCCTTAATCCAACTCCAGAAATTCGCC	300		
DB	241	ACTGTTAGAGGCTGAGTGACAGGCTCTTACAGTGGCCCTTAATCCAACTCCAGAAATTCGCC	300		
QY	301	AACGGAACCTTTGAGATTTATGCAATTCGAAGTGAAGGACAGGAAATTCATTCCT	360		
DB	301	AACGGAACCTTTGAGATTTATGCAATTCGAAGTGAAGGACAGGAAATTCATTCCT	360		
QY	361	CTTAATGTACATGGATGGCAAGATGATTCGAGCTCTCTTCCAGTCCGATGAGATG	420		
DB	361	CTTAATGTACATGGATGGCAAGATGATTCGAGCTCTCTTCCAGTCCGATGAGATG	420		
QY	421	GAGATGCCCTTGTCAATGAAGGGCCNCCTGTTGTCAATTCGAGCTTACACAAAGAAAAA	480		
DB	421	GAGATGCCCTTGTCAATGAAGGGCCNCCTGTTGTCAATTCGAGCTTACACAAAGAAAAA	480		
QY	481	ATGTCAATCCGAATCGAGGGGAATATGCTTGGATTCGATTCGAGCTTACAGGCTTAC	540		
DB	481	ATGTCAATCCGAATCGAGGGGAATATGCTTGGATTCGATTCGAGCTTACAGGCTTAC	540		
QY	541	ACACATTCAGAGACCTTTAATAAATGCTTAAATGCAACACCGGCTTACCTCTGTGAA	600		
DB	541	ACACATTCAGAGACCTTTAATAAATGCTTAAATGCAACACCGGCTTACCTCTGTGAA	600		
QY	601	CCAGCAGTCTTTCGGGTGAAGCAGAGTATCTCAGTCCGCTTGTATGATAAAGTCAAGTGCA	660		
DB	601	CCAGCAGTCTTTCGGGTGAAGCAGAGTATCTCAGTCCGCTTGTATGATAAAGTCAAGTGCA	660		
QY	661	ACAGAACCTTCCCAAGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATGTGG	720		
DB	661	ACAGAACCTTCCCAAGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATGTGG	720		
QY	721	CAGACATTTAGAGTCGCTTTGATGTTGATTCGATCCCATGAGAACACACAAAGTTCTTTC	780		
DB	721	CAGACATTTAGAGTCGCTTTGATGTTGATTCGATCCCATGAGAACACACAAAGTTCTTTC	780		
QY	781	ACTTACGGGTGTAACATGTGCGGAAGAGATTCAAGGAGCCTTGTGTTCTTAAAAATCAC	840		
DB	781	ACTTACGGGTGTAACATGTGCGGAAGAGATTCAAGGAGCCTTGTGTTCTTAAAAATCAC	840		
QY	841	ATGCGGACRCATATATGCGAAATCGGGGCGCAGAACAACTCAGCAAGGCTTGAGAGT	900		
DB	841	ATGCGGACRCATATATGCGAAATCGGGGCGCAGAACAACTCAGCAAGGCTTGAGAGT	900		
QY	901	AGTCAGGACACATCAACGAGTCTGTCAGGTGTCAGGGGCGCAGAGCATCTCTCTCT	960		
DB	901	AGTCAGGACACATCAACGAGTCTGTCAGGTGTCAGGGGCGCAGAGCATCTCTCTCT	960		
QY	961	TGCAAAATCTGATGTTGTGGCTTCTCTTATTTTCAAAATAAAGAAAGTCTAATTCAGCAC	1020		
DB	961	TGCAAAATCTGATGTTGTGGCTTCTCTTATTTTCAAAATAAAGAAAGTCTAATTCAGCAC	1020		
QY	1021	CGCAAGGTGCACACCAAAAAAACTGCTTTTCGGTACCAGCAGCGCAGACACTCTCCA	1080		

Query Match 92.1%; Score 1387.2; DB 4; Length 10365;
Best Local Similarity 97.3%; Pred. No. 0; Mismatches 3; Indels 6; Gaps 6;
Matches 1470; Conservative

RESULT 3
US-08-892-695-9
; Sequence 9, Application US/08892695A
; Patent No. 6808878
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W
; APPLICANT: Collins, Collin
; APPLICANT: Hwang, Soo In
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: GENES FROM THE 20Q13 AMPLICON AND THEIR USES
; FILE REFERENCE: 2500.124US3
; CURRENT APPLICATION NUMBER: US/08/892,695A
; CURRENT FILING DATE: 1997-07-15
; EARLIER APPLICATION NUMBER: 08/785,532
; EARLIER FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/731,499
; EARLIER FILING DATE: 1996-10-16
; EARLIER APPLICATION NUMBER: 08/680,395
; EARLIER FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 10365
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Genomic
; OTHER INFORMATION: Sequence encoding ZABCI
; Patent No. 6808878
US-08-892-695-9

QY 2 CAGGTTGCGGATGACTTCTTCTCAATTGAACACTCATTTCAATGAGCAAGAGC 61
DB |||||||
2424 CAGGTTGCTGGGATTTGACTTTCTGTCTCAATTGAAACACTCATTTCAATGAGCAAGAGC 2483
QY 62 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTTGGGAACCCCTGTATGCTTTGT 121
DB |||||||
2484 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTTGGGAACCCCTGTATGCTTTGT 2543
QY 122 TTGTGAAAGAACGATGACACCATCATCTGAGCTTCTTAAAGTTTGAAGAGTTAGAGG 181
DB |||||||
2544 TTGTGAAAGAACGATGACACCATCATCTGAGCTTCTTAAAGTTTGAAGAGTTAGAGG 2603
QY 182 ACTATACACTTTCTTTTGAACCTTTTATAATAATATTTGCTCTGG-TTTTGGAAACCCAGG 240
DB |||||||
2604 ACTATACACTTTCTTTTGAACCTTTTATAATAATATTTGCTCTGGTTTITGGAAACCCAGG 2663
QY 241 ACTGTTAGA-GGGTGAGTGACAGGCTTTAC-AGTGGCCCTTAATCCAACTCCAGAAATTTGC 298
DB |||||||
2664 GCTGTTAGAGGGTGAGTGACAAGTCTTACAAGTGGCCCTTATTCCACTCCAGAAATTTGC 2723
QY 299 CCAACGGAACTTTGAGATTATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 358
DB |||||||
2724 CCAACGGAACTTTGAGATTATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 2783
QY 359 CTCCTTAATGTACATGAT-GGCCAAGAGTGATTGGCAGCTCTCTTGCCAGTCCGATGGAG 417
DB |||||||
2784 CTCCTTAATGTACATGATGGCCGAGAGTGATTGGCAGCTCTCTTGCCAGTCCGATGGAG 2843
QY 418 ATGGA-GATGCCCTTGTCAATGAAGGGCCNCTGTTGTCAATTCGAGCTACACAAAGAA 476
DB |||||||
2844 ATGAGGATGCTTGTCAATGAAGGGCCGCTGTTGTTCCATTCGAGCTACACAAAGAA 2903
QY 477 AAAATGTCAATCCGAATCGAGGGGAATATGCCCTTTGGATTGCATGTTCTGCGAGCCAGC 536
DB |||||||
2904 AAAATGTG-ATCCAAATCGAGGGGTATATGCCCTTTGGATTGCATGTTCTGCGAGCCAGC 2962
QY 537 CTTACACATTCAGAAAGACTTTAATAACATGCTTTAATGCAACACGGCCCTACCCCTCTG 596
DB |||||||
2963 CTTACACATTCAGAAAGACTTTAATAACATGCTTTAATGCAACACGGCCCTACCCCTCTG 3022
QY 597 TGAACACGAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGT 656
DB |||||||
3023 TGAACACGAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGT 3082
QY 657 GCGAACAGAACTTCCACAGGAAAGAAATTCAGGAAATGAATTTAGCTGTGAGGTATG 716
DB |||||||
3083 GCGAACAGAACTTCCACAGGAAAGAAATTCAGGAAATGAATTTAGCTGTGAGGTATG 3142
QY 717 TGGGAGACATTTAGAGTCGCTTTTGTGTTGAGATCCACATGAGAACACACAAAGATTC 776
DB |||||||
3143 TGGGAGACATTTAGAGTCGCTTTTGTGTTGAGATCCACATGAGAACACACAAAGATTC 3202
QY 777 TTTCACTTACGGGTGAACATGTGCGGAAGAGATTTCAAGGAGCCTTGCTTTCTTAAAAA 836
DB |||||||
3203 TTTCACTTACGGGTGAACATGTGCGGAAGAGATTTCAAGGAGCCTTGCTTTCTTAAAAA 3262
QY 837 TCACATGCGGACRCAATATGCAATTCGGGGCCAGAGCAAACTGACGAGGCTTTCGA 896
DB |||||||
3263 TCACATGCGGACACATAATGCGCAATTCGGGGCCAGAGCAAACTGACGAGGCTTTCGA 3322
QY 897 GAGTAGTCCAGCAACGATCAACAGAGTCTGTGAGTGCACGCGGCGGAGAGCATCTCTCT 956
DB |||||||
3323 GAGTAGTCCAGCAACGATCAACAGAGTCTGTGAGTGCACGCGGCGGAGAGCATCTCTCT 3382
QY 957 TCCTTGAACAAATCTGCATGTTTGTGGCTTCTATTTCAAATAAAGAAAGTCTAATTTGA 1016
DB |||||||
3383 TCCTTGAACAAATCTGCATGTTTGTGGCTTCTATTTCAAATAAAGAAAGTCTAATTTGA 3442
QY 1017 GCACGCAAGGTGCACCAACCAAACTGCTTTTCGTTACCGACGCGGCGGAGAGCATC 1076
DB |||||||
3443 GCACGCAAGGTGCACCAACCAAACTGCTTTTCGTTACCGACGCGGCGGAGAGCATC 3502
QY 1077 TCCACAGGAGGAATGCGGCTCTCGAGGGAGGACTTCTCTGCAAGTTGTTTCAACTTGAGACC 1136

DB 3503 TCCACAGAGGAATGCCGTCTCTCGAGGAGGACTTCTCTGAGTTGTTCAACTTGAGACC 3562
QY 1137 AAAATCTCACCTGAAACGGGGAAGAACCTGTTCAGATGCAATCCCTCAGCTCGATCCGTT 1196
DB |||||||
3563 AAAATCTCACCTGAAACGGGGAAGAACCTGTTCAGATGCAATCCCTCAGCTCGATCCGTT 3622
QY 1197 CACCACCTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGAGTTGCCATTTGCCAAGAGT 1256
DB |||||||
3623 CACCACCTTCCAGGCTTGGCAGCTGGCTACCAAGGAAGAAAGTTGCCATTTGCCAAGAGT 3682
QY 1257 GAAGGAATTTGGGGAAGGAAGGAGCACCACAAACGACGATTTCGAGTTCGAGAAAGAGCT 1316
DB |||||||
3683 GAAGGAATTCGGGCGAGGAAGGAGCACCACAAACGACGATTTCGAGTTCGAGAAAGAGCT 3742
QY 1317 TGGAGAAAACAAATGAAGAACCAATTTGTCAGGCTCTCGCAAGAGAAAGAGTGCACAA 1376
DB |||||||
3743 TGGAGAAAACAAATGAAGGCGAGTTGTGCAAGGCTCTCGCAAGAGAAAGAGTGCACAA 3802
QY 1377 CTCACAGGCGAAGGGCCCTCGTGGAGCGGATCCCAAGTTACCCAGTAGCAAGAGNA 1436
DB |||||||
3803 CTCACAGGCGAAGGGCCCTCGTGGAGCGGATCCCAAGTTACCCAGTAGCAAGAGNA 3862
QY 1437 GCCCACTCACTGCTCCGAGTGGCGCAAGCTTTTCAGAACCTTACCAACAGCTGCTTGCA 1496
DB |||||||
3863 GCCCACTCACTGCTCCGAGTGGCGCAAGCTTTTCAGAACCTTACCAACAGCTGCTTGCA 3922
QY 1497 CTCAGGGTCC 1507
DB |||||||
3923 CTCAGGGTCC 3933

RESULT 4

US-09-949-016-12604
; Sequence 12604, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12604
; LENGTH: 20022
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12604

Query Match 92.1%; Score 1387.2; DB 4; Length 20022;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1470; Conservative 3; Mismatches 32; Indels 6; Gaps 6;

QY 2 CAGGTTGCTGGGATTTGACTTCTTCTCAATTGAAACACTCATTTCAATGAGCAAGAGC 61
DB |||||||
1951 CAGGTTGCTGGGATTTGACTTCTTCTCAATTGAAACACTCATTTCAATGAGCAAGAGC 2010
QY 62 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTTGGGAACCCCTGTATGCTTTGT 121
DB |||||||
2011 ACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTTGGGAACCCCTGTATGCTTTGT 2070
QY 122 TTGTGAAAGAACCGAGTGACACCATCACTGAGCTTCTTAAAGTTTCGAAGAGTTAGAGG 181
DB |||||||
2071 TTGTGAAAGAACCGAGTGACACCATCACTGAGCTTCTTAAAGTTTCGAAGAGTTAGAGG 2130

QY 182 ACTATACACTTTCTTTTGAACCTTTTATAATAAATTTGCTCTGG-TTTTGAACCCAGG 240
DB 2131 ACTATACACTTTCTTTTGAACCTTTTATAATAAATTTGCTCTGGTTTGTGAACCCAGG 2190
QY 241 ACTGTTAGA-GGGTAGTCAGAGCTCTTAC-AGTGGCCCTTAATCCAACTCCAGAAATTCG 298
DB 2191 GCTGTTAGAGGGGTGAGTGACAAGCTTTTACAAGTGGCCCTTATTTCCAACTCCAGAAATTCG 2250
QY 299 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTCAGAGGAAACATGCAACTCAATCC 358
DB 2251 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTCAGAGGAAACATGCAACTCAATCC 2310
QY 359 CTCCTTAATGATACATGAT-GGCCAAGAGTATTGGCAGCTCTCTTGCAGTCCGATGGAG 417
DB 2311 CTCCTTAATGATACATGATGGCCGAGAGTGAATGGCAGCTCTCTTGGCAGTCCGATGGAG 2370
QY 418 ATGGA-GATGCCCTTGTCAATGAAGGGCCNCTGTTGTCATTCAGTCCGAGCTACACAAGAA 476
DB 2371 ATGGAGATGCTTGTCAATGAAGGGACCGCTGTTGTTCCATTCGAGCTACACAAGAA 2430
QY 477 AAAAAATGTCATCCGAATCCAGGGGAATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 536
DB 2431 AAAAAATGTC-ATCCAAATCCAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 2489
QY 537 CTTACACATTCAGAGACCTTAATAAACAATGTCTTAATGCAACACCGCCCTACCCCTCTG 596
DB 2490 CTTACACATTCAGAGACCTTAATAAACAATGTCTTAATGCAACACCGCCCTACCCCTCTG 2549
QY 597 TGAACCCAGAGTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCMAGT 656
DB 2550 TGAACCCAGAGTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCMAGT 2609
QY 657 GCGAACAGAACTCTCCAGGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATG 716
DB 2610 GCGAACAGAACTCTCCAGGAAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATG 2669
QY 717 TGGGAGACATTTAGAGTCGCTTTTGAATGTTGAGATCCACATGAGAAACACAAAGATTC 776
DB 2670 TGGGAGACATTTAGAGTCGCTTTTGAATGTTGAGATCCACATGAGAAACACAAAGATTC 2729
QY 777 TTTTCACTTACGGGTGTAACATGTCGGGAGAGATTCAGAGGAGCTTGGTTCTTAAAAA 836
DB 2730 TTTTCACTTACGGGTGTAACATGTCGGGAGAGATTCAGAGGAGCTTGGTTCTTAAAAA 2789
QY 837 TCACATGCGGACRCATAATGCGAAATTCGGGGCGCAGAACTGCAGCAAGGCTTGA 896
DB 2790 TCACATGCGGACRCATAATGCGAAATTCGGGGCGCAGAACTGCAGCAAGGCTTGA 2849
QY 897 GAGTAGTCAGCAACGATCAACGAGGTCTGTCAGGTGCAACGCGGCGGAGAGCATCTCCTC 956
DB 2850 GAGTAGTCAGCAACGATCAACGAGGTCTGTCAGGTGCAACGCGGCGGAGAGCATCTCCTC 2909
QY 957 TCCCTTGAATAATCTGATGTTTGGCTTCTCTATTTCCAAATGAAGAAGTCTAATGA 1016
DB 2910 TCCCTTGAATAATCTGATGTTTGGCTTCTCTATTTCCAAATGAAGAAGTCTAATGA 2969
QY 1017 GCACCGCAAGGTGCAACCAAAAAAATGCTTTCCGTTACAGCAGCGCGCAGCAGACTC 1076
DB 2970 GCACCGCAAGGTGCAACCAAAAAAATGCTTTCCGTTACAGCAGCGCGCAGCAGACTC 3029
QY 1077 TCCCAAGGAGGAATGCGCTCTCGAGGAGGACTTCTCGCAGTGTGTTCAACTTGAGACC 1136
DB 3030 TCCCAAGGAGGAATGCGCTCTCGAGGAGGACTTCTCGCAGTGTGTTCAACTTGAGACC 3089
QY 1137 AAAATCTCACCTGAAACGGGGAAGAGCTGTGATGTCATCCCTCAGCTCGATCCGTT 1196
DB 3090 AAAATCTCACCTGAAACGGGGAAGAGCTGTGATGTCATCCCTCAGCTCGATCCGTT 3149
QY 1197 CACCACCTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGAGTTCCTATTTGCCAAGAAGT 1256
DB 3150 CACCACCTTCCAGGCTTGGCAGCTGGCTACCAAGGAAGAGTTCCTATTTGCCAAGAAGT 3209
QY 1257 GAAGGAATTTGGGCAAGAGGGAGCACCGCAACCGACGATTCGAGTTCGAGAGAGGAGCT 1316

DB 3210 GAAGGAATTCGGGCAAGAGGGAGCACCAACGACGATTCGAGTTCGAGAGGAGCT 3269
QY 1317 TCGAGAAACAAATAAAGAACCAATTTGTGAGGCTCTTCGCAAGAGAAAGAGAGTGAACA 1376
DB 3270 TCGAGAAACAAATAAAGGCGATTTGTGAGGCTCTTCGCAAGAGAAAGAGAGTGAACA 3329
QY 1377 CTCCACGCGCAAGCGCCCTCCGTCGACGCGGATCCCAAGTTACCCAGTAGCAAGAGAA 1436
DB 3330 CTCCACGCGCAAGCGCCCTCCGTCGAGCGGATCCCAAGTTACCCAGTAGCAAGAGAA 3389
QY 1437 GCCCACTCACTCTCCGAGTCCGCAAGCTTTTCAGAACTTACCACCACTGGTCTTGA 1496
DB 3390 GCCCACTCACTCTCCGAGTCCGCAAGCTTTTCAGAACTTACCACCACTGGTCTTGA 3449
QY 1497 CTCCAGGGTCC 1507
DB 3450 CTCCAGGGTCC 3460

RESULT 5
US-09-949-016-16004
; Sequence 16004, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE OF INVENTION: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16004
; LENGTH: 20023
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16004

Query Match 92.1%; Score 1387.2; DB 4; Length 20023;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1470; Conservative 3; Mismatches 32; Indels 6; Gaps 6;
QY 2 CAGGTTGCTGGGATTGACTTCTTGTCTCAATTGAAACACTCATTCAATGAGACAAAGAG 61
DB 1951 CAGGTTGCTGGGATTGACTTCTTGTCTCAATTGAAACACTCATTCAATGAGACAAAGAG 2010
QY 62 ACTAATGCTTTGCTGATTCATATTGATCGAGCATTTGGAAACCTTGTATGCCTTGT 121
DB 2011 ACTAATGCTTTGCTGATTCATATTGATCGAGCATTTGGAAACCTTGTATGCCTTGT 2070
QY 122 TTGTGGAAGAACACAGTGCACCATCTGAGCTTCTTAAATATTTGCTCTGG-TTTTGAACCCAGG 181
DB 2071 TTGTGGAAGAACACAGTGCACCATCTGAGCTTCTTAAATATTTGCTCTGGTTTGTGAACCCAGG 2130
QY 182 ACTATACACTTTCTTTTGAACCTTTTATAATAAATTTGCTCTGG-TTTTGAACCCAGG 240
DB 2131 ACTATACACTTTCTTTTGAACCTTTTATAATAAATTTGCTCTGGTTTGTGAACCCAGG 2190
QY 241 ACTGTTAGA-GGGTAGTCAGAGCTTTCAC-AGTGGCCCTTAATCCAACTCCAGAAATTCG 298
DB 2191 GCTGTTAGAGGGGTGAGTGACAAGCTTTTACAAGTGGCCCTTATTTCCAACTCCAGAAATTCG 2250
QY 299 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTCAGAGGAAACATGCAACTCAATCC 358
DB 2251 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTCAGAGGAAACATGCAACTCAATCC 2310

QY 359 CTCCTTAATGTCATCGAT-GGCCAAGAGTGATTGGCAGCTCTCTTCCAGTCCGATGGAG 417
DB |||||
QY 2311 CTCCTTAATGTCATCGATGGCCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAG 2370
DB |||||
QY 418 ATGGA-GATGCCCTTGTCAATGAAGAGGCCNCTGTGTCAATTCGCGAGCTACACAAAGAA 476
DB |||||
QY 2371 ATGGAGGATGCCCTTGTCAATGAAGAGGACCGCTGTGTTCATTCGAGCTACACAAGAA 2430
DB |||||
QY 477 AAAATGTCAATCCGAATCGAGGGAATATGCCCTTGGATTCGATTCGTCAGCGACAC 536
DB |||||
QY 2431 AAAATGTG-ATCCAAATCGAGGGGTATATGCCCTTGGATTCGATTCGTCAGCGACAC 2489
DB |||||
QY 537 CTCACACATTCAGAGACCTTAATAAACAATGCTTTAATGAACAACCGSCCTACCCCTCTG 596
DB |||||
QY 2490 CTTTACACATTCAGAGACCTTAATAAACAATGCTTTAATGAACAACCGSCCTACCCCTCTG 2549
DB |||||
QY 597 TGAACACGAGTTCCTCGGGTTGAAGCAGAGTATCTCAGTCGCTTGATAAAGTCAAGT 656
DB |||||
QY 2550 TGAACACGAGTTCCTCGGGTTGAAGCAGAGTATCTCAGTCGCTTGATAAAGTCAAGT 2609
DB |||||
QY 657 GCGAACAGAACTCCCAAGGAAAGAAATGCNAGGAAATGAATTTAGCTGTGAGGTATG 716
DB |||||
QY 2610 GCGAACAGAACTCCCAAGGAAAGAAATGCNAGGAAATGAATTTAGCTGTGAGGTATG 2669
DB |||||
QY 717 TGGGCGAGACATTTAGAGTCGCTTTTGTGTTGATCCACATGAGAACACACAAAGATTC 776
DB |||||
QY 2670 TGGGCGAGACATTTAGAGTCGCTTTTGTGTTGATCCACATGAGAACACACAAAGATTC 2729
DB |||||
QY 777 TTTTCACTTACGGGTGTAAATGTCGGAAGAGATTCGAAGGAGCCTTGCTTTCTTAAAAA 836
DB |||||
QY 2730 TTTTCACTTACGGGTGTAAATGTCGGAAGAGATTCGAAGGAGCCTTGCTTTCTTAAAAA 2789
DB |||||
QY 837 TCACATGCGGACRATATATGCAATTCGGGGCCAGAGCAAACTGCAGCAAGGCTTGA 896
DB |||||
QY 2790 TCACATGCGGACRATATATGCAATTCGGGGCCAGAGCAAACTGCAGCAAGGCTTGA 2849
DB |||||
QY 897 GAGTAGTCAGCAACGATCAACAGAGTCGTCCAGGTGCACGGCGCGAGAGCATCTCCTC 956
DB |||||
QY 2850 GAGTAGTCAGCAACGATCAACAGAGTCGTCCAGGTGCACGGCGCGAGAGCATCTCCTC 2909
DB |||||
QY 957 TCCTTGTCAAAATCTGCATGGTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTTGA 1016
DB |||||
QY 2910 TCCTTGTCAAAATCTGCATGGTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTTGA 2969
DB |||||
QY 1017 GCACGCGAAGTGCACACCAAAACTGCTTCGGTACCGAGCAGCGCGCAGACAGATC 1076
DB |||||
QY 2970 GCACGCGAAGTGCACACCAAAACTGCTTCGGTACCGAGCAGCGCGCAGACAGATC 3029
DB |||||
QY 1077 TCCCAAGGAGGAATGCCGTCTCGAGGAGGACCTTCCTCGAGTTGTTCAACTTGAGACC 1136
DB |||||
QY 3030 TCCCAAGGAGGAATGCCGTCTCGAGGAGGACCTTCCTCGAGTTGTTCAACTTGAGACC 3089
DB |||||
QY 1137 AAAATCTCACCTGAAACGGGGAAGAGCCTGTTCAGATGCAATCCCTCAGCTCGATCCGTT 1196
DB |||||
QY 3090 AAAATCTCACCTGAAACGGGGAAGAGCCTGTTCAGATGCAATCCCTCAGCTCGATCCGTT 3149
DB |||||
QY 1197 CACCACTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGTGTGCCATTTGCCRAGAGT 1256
DB |||||
QY 3150 CACCACTTCCAGGCTTGGCAGCTGGCTACCAAGGAAGAAAGTTGCCATTTGCCAAGAGT 3209
DB |||||
QY 1257 GAAGGAATTTGGGCAAGAGGAGCAGCAACGACGATTCGAGTTCGAGAGAGGAGCT 1316
DB |||||
QY 3210 GAAGGAATTCGGGCAAGAGGAGCAGCAACGACGATTCGAGTTCGAGAGAGGAGCT 3269
DB |||||
QY 1317 TGGAGAAAACAAATAGAAACCAATTTGTGAGGCTCTTCGCAAGAGAAAGAGAGTGCAGAA 1376
DB |||||
QY 3270 TGGAGAAAACAAATAGAGGAGTGTGTCAGGCTCTTCGCAAGAGAAAGAGAGTGCAGAA 3329
DB |||||
QY 1377 CTCACCGGAGAGCGCCTCGTGGACCGGATCCCAAGTTACCAGTACGAGGAGAA 1436
DB |||||
QY 3330 CTCACCGGAGAGCGCCTCGTGGACCGGATCCCAAGTTACCAGTACGAGGAGAA 3389
DB |||||
QY 1437 GCCCACTCACTGCTCCGAGTGGCGAAAGCTTTTCAGAACTTACCACCGAGCTGGTCTTGCA 1496
DB |||||

DB 3390 GCCCACTCACTGCTCCGAGTGGCGCAAGCTTTTCAGAACCTTACCACGAGCTGGTCTTGCA 3449
QY 1497 CTCACGAGGTCC 1507
DB 3450 CTCACGAGGTCC 3460

RESULT 6
US-09-560-594-3
; Sequence 3, Application US/09560594
; Patent No. 6242590
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowser
; TITLE OF INVENTION: ANTISENSE MODULATION OF ZINC FINGER PROTEIN-217 EXPRESSION
; FILE REFERENCE: RTS-0144
; CURRENT APPLICATION NUMBER: US/09/560,594
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 3
; LENGTH: 5632
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (272)...(3418)
US-09-560-594-3

Query Match 88.7%; Score 1337.2; DB 3; Length 5632;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1420; Conservative 3; Mismatches 32; Indels 6; Gaps 6;

QY 52 GACAAAGAGCACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTCGGAAACCCCTG 111
DB 1 GACAAAGAGCACTAATGCTTTGTGCTGATTCATATTTGAATCGAGGCATTCGGAAACCCCTG 60

QY 112 TATGCTTTGTTGTGGAAGAACCAAGTCACACCATCACTGAGCTTCCTTAAAGTTTCGAAG 171
DB 61 TATGCTTTGTTGTGGAAGAACCAAGTCACACCATCACTGAGCTTCCTTAAAGTTTCGAAG 120

QY 172 AAGTTAGAGGACTATACACTTTTCTTTTGAATTTTATATAAATTTTCTCTGG-TTTT 230
DB 121 AAGTTAGAGGACTATACACTTTTCTTTTGAATTTTATATAAATTTTCTCTGGTTT 180

QY 231 GGAACCCAGGAGCTGTAGA-GGGTGAAGTGAAGTCTTAC-AGTGGCCCTTAATCCAACTC 288
DB 181 GGAACCCAGGAGCTGTAGAGGGGTGAGTGACAAGTCTTTACAAGTGGCCTTATTTCCAACTC 240

QY 289 CAGAAATGGCCCAACCGGAATTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 348
DB 241 CAGAAATGGCCCAACCGGAATTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300

QY 349 AACTCAATCCCTCTTAATGTACATGGAT-GGCACAGAGTGTGGCAGCTCTCTTGGCAG 407
DB 301 AACTCAATCCCTCTTAATGTACATGGATGGGCGCAGAGTGTGGCAGCTCTCTTGGCAG 360

QY 408 TCCGATGAGATGA-GATGCTTGTCAATGAAGGGCCCTGTTGTCAATTCAGAGCT 466
DB 361 TCCGATGAGATGAGGATGCTTGTCAATGAAGGGACCGCTGTGTGTTCATTCAGAGCT 420

QY 467 ACACAAAGAAAAAATGTCATTCGAAATCGAGGGGAATATGCCCTTGGATTGCATGTTCT 526
DB 421 TACACAGAAAAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCT 479

QY 527 GCAGCAGACCTTCAACATTCAGAAAGCCTTATAAATGATGCTTCTTAATGCAACACCGGC 586
DB 480 GCAGCAGACCTTCAACATTCAGAAAGCCTTATAAATGATGCTTCTTAATGCAACACCGGC 539

QY 587 CTACCTCTGTGGAACCAAGCTTCTTCGGGTGGAAGCAGAGTATCTCAGTCCGCTTGATA 646
DB 540 CTACCTCTGTGGAACCAAGCTTCTTCGGGTGGAAGCAGAGTATCTCAGTCCGCTTGATA 599

QY 647 AAAGTCAAGTGCAGAACAGACCTCCCAAGGAAAAAGATTGCAAGGAAAAATGAATTTAGCT 706
DB |||||

Db 600 AAAGTCAAGTGGGAAACAGAACCTCCCAAGGAAAGAAATTCGAAGGAAATGAATTTAGCT 659
Qy 707 GTGAGGTATGTGGGACAGACATTTAGAGTGCCTTTTGTGATGTTGAGATCCACATGAGAACAC 766
Db 660 GTGAGGTATGTGGGACAGACATTTAGAGTGCCTTTTGTGATGTTGAGATCCACATGAGAACAC 719
Qy 767 ACAAGATCTTTTCACTTACCGGTGTAACATGTGCGGAAAGAGATTCGAAGGAGCTTTGT 826
Db 720 ACAAGATCTTTTCACTTACCGGTGTAACATGTGCGGAAAGAGATTCGAAGGAGCTTTGT 779
Qy 827 TTCTTAAATATCATGCGGACBCATAATGCGAAATCGGGGCGAGAACCAATGTCAGC 886
Db 780 TTCTTAAATATCATGCGGACBCATAATGCGAAATCGGGGCGAGAACCAATGTCAGC 839
Qy 887 AAGGCTTGAGAGTAGTCCAGCAACGATCAACGAGGTGTCAGGTGTCAGCGGCGGAG 946
Db 840 AAGGCTTGAGAGTAGTCCAGCAACGATCAACGAGGTGTCAGGTGTCAGCGGCGGAG 899
Qy 947 GCATCTCTCTCTCTGCAAAATCTGCATGTTTGTGGCTTCTTATTTCAAAATAAGAA 1006
Db 900 GCATCTCTCTCTCTGCAAAATCTGCATGTTTGTGGCTTCTTATTTCAAAATAAGAA 959
Qy 1007 GTCTAATTTAGCACCGCAGAGTGCCACCAAAAAAATCTGTTTCTGATACAGAGCGGC 1066
Db 960 GTCTAATTTAGCACCGCAGAGTGCCACCAAAAAAATCTGTTTCTGATACAGAGCGGC 1019
Qy 1067 AGACAGACTCTCCACAAGGAGGAAATGCGCTCTCGAGGAGGACTTCTGCGAGTTGTTCA 1126
Db 1020 AGACAGACTCTCCACAAGGAGGAAATGCGCTCTCGAGGAGGACTTCTGCGAGTTGTTCA 1079
Qy 1127 ACTTGAGACCAAAATCTCAACCTGAAACGGGGAAGAGCTGTGAGATGATCTCCCTCAGC 1186
Db 1080 ACTTGAGACCAAAATCTCAACCTGAAACGGGGAAGAGCTGTGAGATGATCTCCCTCAGC 1139
Qy 1187 TCGATCCGTTTACACCTTCCAGGCTTGGCACTGTGCTACCAAGGAAAGTTGCCATTT 1246
Db 1140 TCGATCCGTTTACACCTTCCAGGCTTGGCACTGTGCTACCAAGGAAAGTTGCCATTT 1199
Qy 1247 GCCAAGAGTGAAGAAATTTGGGGAAGAGGAGGACCCACACGACGATTCGAGTTCCG 1306
Db 1200 GCCAAGAGTGAAGAAATTTGGGGAAGAGGAGGACCCACACGACGATTCGAGTTCCG 1259
Qy 1307 AGAAGAGCTTGGAGAAACAAATAAGAACCTTGTGAGGCTCTGCGAAGAGAAAGAGA 1366
Db 1260 AGAAGAGCTTGGAGAAACAAATAAGAACCTTGTGAGGCTCTGCGAAGAGAAAGAGA 1319
Qy 1367 AGTGCAACACTCCACGCGGAGGCGCTTCCGTGGACGCGGATCCCAAGTTACCCAGTA 1426
Db 1320 AGTGCAACACTCCACGCGGAGGCGCTTCCGTGGACGCGGATCCCAAGTTACCCAGTA 1379
Qy 1427 GCAAGGAGAGCCCTCACTCTCCGAGTGGGCAAGCTTTCGAACTTACCCAGC 1486
Db 1380 GCAAGGAGAGCCCTCACTCTCCGAGTGGGCAAGCTTTCGAACTTACCCAGC 1439
Qy 1487 TGGTCTTGCACTCCAGGGTCC 1507
Db 1440 TGGTCTTGCACTCCAGGGTCC 1460

RESULT 7
US-09-949-016-862
; Sequence 862, Application US/09949016
; Patent No. 6812319
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 862
; LENGTH: 5632
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-862

Query Match 88.7%; Score 1337.2; DB 4; Length 5632;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1420; Conservative 3; Mismatches 32; Indels 6; Gaps 6;

Qy 52 GACAAAGAGACATAATGCTTGTGCTGATTTGATATTTGAATCGAGCATTTGGAAACCTTG 111
Db 1 GACAAAGAGACATAATGCTTGTGCTGATTTGATATTTGAATCGAGCATTTGGAAACCTTG 60
Qy 112 TATGCTTGTGCTGAAAGAACCAAGTGCACCACTCAGCTTCTTAAAAAGTTTCAAAG 171
Db 61 TATGCTTGTGCTGAAAGAACCAAGTGCACCACTCAGCTTCTTAAAAAGTTTCAAAG 120
Qy 172 AAGTTAGAGGACTATACACTTTCTTTGAACTTTTATATAAATATTTGCTCTGG-TTTT 230
Db 121 AAGTTAGAGGACTATACACTTTCTTTGAACTTTTATATAAATATTTGCTCTGGTTTT 180
Qy 231 GGAACCCAGGAGTGTAG- GGTGAGTGCAGAGTCTTAC- AGTGGCTTAAATCCAACTC 288
Db 181 GGAACCCAGGAGTGTAGAGGGGTGAGTGACAAAGTCTTACAAGTGGGCTTATTTCAAATC 240
Qy 289 CAGAAATTTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 348
Db 241 CAGAAATTTGCCCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300
Qy 349 AACTCAATCCCTCTTAAATGATACATGAT- GGCCAAAGAGTGTATGGCAGCTCTCTTCCAG 407
Db 301 AACTCAATCCCTCTTAAATGATACATGATGGGCGCAGAAAGTATTTGGCAGCTCTCTTCCAG 360
Qy 408 TCGATGAGAGTGA- GATGCTTGTCAATGAAAGGCGCCCTGTGTGTCATTTCCAGCT 466
Db 361 TCGATGAGAGTGAAGGATGCTTGTCAATGAAAGGCGCGCTGTGTGTCATTTCCAGCT 420
Qy 467 ACACAAAGAAAAAATGTCAATCCGAATCGAGGGGAAATATGCGCTTTGGATTTGATTTCT 526
Db 421 TACACAGAAAAAATGTC- ATCCAAATCGAGGGGTATATGCGCTTTGGATTTGATTTCT 479
Qy 527 GCAGCCAGACTTTCACATTCAGAAAGCTTTAATAAATGCTTTAATGCAACACCGGC 586
Db 480 GCAGCCAGACTTTCACATTCAGAAAGCTTTAATAAATGCTTTAATGCAACACCGGC 539
Qy 587 CTACCTCTGTGAACCGCAGTCTTTCGGTGTGAAGAGAGTATCTCAGTCCGCTTGATA 646
Db 540 CTACCTCTGTGAACCGCAGTCTTTCGGTGTGAAGAGAGTATCTCAGTCCGCTTGATA 599
Qy 647 AAAGTCAAGTGCAGAACAGAACTCCCAAGAAAAAGAAATGCAAGGAAATGAATTTAGCT 706
Db 600 AAAGTCAAGTGCAGAACAGAACTCCCAAGAAAAAGAAATGCAAGGAAATGAATTTAGCT 659
Qy 707 GTGAGGTATGTGGGAGACATTTAGAGTGCCTTTTGTGATGTTGAGATCCACATGAGAACAC 766
Db 660 GTGAGGTATGTGGGAGACATTTAGAGTGCCTTTTGTGATGTTGAGATCCACATGAGAACAC 719
Qy 767 ACAAGATCTTTTCACTTACGGGTGTAACTGTGCGGAGAGAGATTCGAAGGAGCTTTGT 826
Db 720 ACAAGATCTTTTCACTTACGGGTGTAACTGTGCGGAGAGAGATTCGAAGGAGCTTTGT 779
Qy 827 TTCTTAAAAATCAGATCGGACRCATATGCGAAATTCGGGGGCGCAGAACCAATTCGACG 886
Db 780 TTCTTAAAAATCAGATCGGACRCATATGCGAAATTCGGGGGCGCAGAACCAATTCGACG 839
Qy 887 AAGGCTTGAGAGTAGTCCAGCAACGATCAACGAGGTGTCAGGTGTCAGCGGCGGAG 946

Db 840 AAGGCTTGGAGAGTAGTCCAGCAACGATCAAGAGGTGTCAGGTGACGCGGCGGAGA 899
Qy 947 GCATCTCTCTCTCTCTGCAAAATCTGCATGTTGGTTCCTATTCCAAATAAGAAA 1006
Db 900 GCATCTCTCTCTCTCTGCAAAATCTGCATGTTGGTTCCTATTCCAAATAAGAAA 959
Qy 1007 GTCTAATTTGAGCACCAGGAGGTGCACACCAAAATCTGTTCCGTACCGAGCGCGC 1066
Db 960 GTCTAATTTGAGCACCAGGAGGTGCACACCAAAATCTGTTCCGTACCGAGCGCGC 1019
Qy 1067 AGACAGACTCTCTCCAAAGGAGGAATGCGCTCTCGAGGGAGGACTTCCTGCAAGTTGTTCA 1126
Db 1020 AGACAGACTCTCTCCAAAGGAGGAATGCGCTCTCGAGGGAGGACTTCCTGCAAGTTGTTCA 1079
Qy 1127 ACTTGAGACCAAAATCTCAACCTGAAACGGGGAAGAACCTGTTCAGATGCAATCCCTCAGC 1186
Db 1080 ACTTGAGACCAAAATCTCAACCTGAAACGGGGAAGAACCTGTTCAGATGCAATCCCTCAGC 1139
Qy 1187 TCGATCCGTTTCAACACCTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGTTGCCATTT 1246
Db 1140 TCGATCCGTTTCAACACCTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGTTGCCATTT 1199
Qy 1247 GCCAAGAGTGAAGGAATTTGGGGAAGAACGGGAGCACCGAACCAAGATTTCGAGTTCCG 1306
Db 1200 GCCAAGAGTGAAGGAATTTGGGGAAGAACGGGAGCACCGAACCAAGATTTCGAGTTCCG 1259
Qy 1307 AGAAGGAGCTTGGAAGAACAAATAAGAACCTATGTCAGGCTCTTCGCAAGAGAAAGAGA 1366
Db 1260 AGAAGGAGCTTGGAAGAACAAATAAGAACCTATGTCAGGCTCTTCGCAAGAGAAAGAGA 1319
Qy 1367 AGTCAACACATCCACCGGAGCGCCCTCGTGAGCGCGATCCCAAGTTACCCAGTA 1426
Db 1320 AGTCAACACATCCACCGGAGCGCCCTCGTGAGCGCGATCCCAAGTTACCCAGTA 1379
Qy 1427 GCAAGGAGAGCCCACTCACTGCTCCGAGTGGGCAAGGCTTTCAGAACTTACCACAGC 1486
Db 1380 GCAAGGAGAGCCCACTCACTGCTCCGAGTGGGCAAGGCTTTCAGAACTTACCACAGC 1439
Qy 1487 TGGTCTTGCACTCCAGGGTCC 1507
Db 1440 TGGTCTTGCACTCCAGGGTCC 1460

RESULT 8

US-09-949-016-4262
; Sequence 4262, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4262
; LENGTH: 5632
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4262

Query Match 88.7%; Score 1337.2; DB 4; Length 5632;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1420; Conservative 3; Mismatches 32; Indels 6; Gaps 6;

Qy 52 GACAAAGAGCACTAATGCTTTGCTGATTCATATTTGAATCGAGGCAATTTGGAAACCCCTG 111
Db 1 GACAAAGAGCACTAATGCTTTGCTGATTCATATTTGAATCGAGGCAATTTGGAAACCCCTG 60
Qy 112 TATGCTTTGTTTGTGGAAAGAACCAAGTGCACCACTCACTGAGCTTCTTAAAGTTTCGAAG 171
Db 61 TATGCTTTGTTTGTGGAAAGAACCAAGTGCACCACTCACTGAGCTTCTTAAAGTTTCGAAG 120
Qy 172 AAGTTAGAGGACTATACACTTTCTTTTGAATTTTATATAATTTGCTCTGG-TTTT 230
Db 121 AAGTTAGAGGACTATACACTTTCTTTTGAATTTTATATAATTTGCTCTGGTTTTT 180
Qy 231 GGAACCCAGGAGTGTTCAG-GGGTGAGTGACAGGTCCTTAC-AGTGGCTTAAATCAACTC 288
Db 181 GGAACCCAGGAGTGTTCAGGGGTGAGTGACAGTCTTACAGTGGCTTATTTCCAACTC 240
Qy 289 CAGAAATTTGCCAAACGGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 348
Db 241 CAGAAATTTGCCAAACGGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300
Qy 349 AACTCAATCCCTCTTAATGTACATGGAT-GGCAAGAGTGATTTGCGAGCTCTTTGCCAG 407
Db 301 AACTCAATCCCTCTTAATGTACATGGATGGGCGCAGAGTGATTTGCGAGCTCTTTGGCAG 360
Qy 408 TCCGATGAGATGGA-GATGCTTGTCAATGAAAGGGCCNCTGTTGTCAATTTCCGAGCT 466
Db 361 TCCGATGAGATGAGGATGCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTTCCGAGC 420
Qy 467 ACACAAGAAAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCT 526
Db 421 TACACAAGAAAAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCT 479
Qy 527 GCAGCAGACCTTCAACATTCAGNAGACCTTAATAACATGTCTTAATGCAACACCCGC 586
Db 480 GCAGCAGACCTTCAACATTCAGNAGACCTTAATAACATGTCTTAATGCAACACCCGC 539
Qy 587 CTACCTCTGTGAACCCAGCAGTCTTTCGGGTGGAAGCAGAGTATCTCAGTCGCTTGATA 646
Db 540 CTACCTCTGTGAACCCAGCAGTCTTTCGGGTGGAAGCAGAGTATCTCAGTCGCTTGATA 599
Qy 647 AAAGTCAAGTGCAGAACAGAACCTCCCAAGGAAAAAGAAATGCAAGAAAAATGAATTAGCT 706
Db 600 AAAGTCAAGTGCAGAACAGAACCTCCCAAGGAAAAAGAAATGCAAGAAAAATGAATTAGCT 659
Qy 707 GTGAGTATGTGGCAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAGAACAC 766
Db 660 GTGAGTATGTGGCAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAGAACAC 719
Qy 767 ACAAGATTTCTTTCACTTACCGGTGTAACATGTGCGGAAGAAAGATTCAAGGAGCCTTGGT 826
Db 720 ACAAGATTTCTTTCACTTACCGGTGTAACATGTGCGGAAGAAAGATTCAAGGAGCCTTGGT 779
Qy 827 TTCTTAAAAATCACATGCGGACRCATATGGAATTCGGGGGCCAGAGAACCTGCAGC 886
Db 780 TTCTTAAAAATCACATGCGGACACATAATGGCAATTCGGGGGCCAGAGAACCTGCAGC 839
Qy 887 AAGGCTTGAGAGTAGTCCAGNAGATCAAGAGTGTCCAGGTGCAAGCGGCGGAGA 946
Db 840 AAGGCTTGAGAGTAGTCCAGNAGATCAAGAGTGTCCAGGTGCAAGCGGCGGAGA 899
Qy 947 GCATCTCTCTCTTTCGCAAAATCTGCATGTTTGTGGCTTCTTATTTCCAAATAAGAAA 1006
Db 900 GCATCTCTCTCTTTCGCAAAATCTGCATGTTTGTGGCTTCTTATTTCCAAATAAGAAA 959
Qy 1007 GTCTAATTTGAGCACCAGGAGGTGCACACCAAAATCTGTTCCGTACCGAGCGCGC 1066
Db 960 GTCTAATTTGAGCACCAGGAGGTGCACACCAAAATCTGTTCCGTACCGAGCGCGC 1019
Qy 1067 AGACAGACTCTCCAAAGGAGGAATGCGCTCTCGAGGGAGGACTTCCTGCAAGTTGTTCA 1126
Db 1020 AGACAGACTCTCCAAAGGAGGAATGCGCTCTCGAGGGAGGACTTCCTGCAAGTTGTTCA 1079
Qy 1127 ACTTGAGACCAAAATCTCAACCTGAAACGGGGAAGAACCTGTTCAGATGCAATCCCTCAGC 1186

Db 1080 ACTTGAGACCAAAATCTCACCCCTGAAACGGGGAAGAGCCTGTGAGATGCATCCCTCAGC 1139
Qy 1187 TCGATCCGTTTACCACCTTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGTGGCATTT 1246
Db 1140 TCGATCCGTTTACCACCTTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGTGGCATTT 1199
Qy 1247 GCCAAGAGTGAAGGAATTTGGGGCAAGAGGAGGACCGCAACGACGATTCGAGTTCCG 1306
Db 1200 GCCAAGAGTGAAGGAATTTGGGGCAAGAGGAGGACCGCAACGACGATTCGAGTTCCG 1259
Qy 1307 AGAAGGAGCTTCGAGAAACAAATAGAACCATTTGTGCAGGCTTCTCGCAAGAGAAGAGA 1366
Db 1260 AGAAGGAGCTTCGAGAAACAAATAGAACCATTTGTGCAGGCTTCTCGCAAGAGAAGAGA 1319
Qy 1367 AGTGAACACACTCCCAACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCAGTA 1426
Db 1320 AGTGCACACACTCCCAACGGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCAGTA 1379
Qy 1427 GCAAGAGAGCCCACTCACTGCTCGAGTGGCGGCAAGCTTTTCAGAACCTTACCACAGC 1486
Db 1380 GCAAGAGAGCCCACTCACTGCTCGAGTGGCGGCAAGCTTTTCAGAACCTTACCACAGC 1439
Qy 1487 TGGTCTTGCACTCCAGGGTCC 1507
Db 1440 TGGTCTTGCACTCCAGGGTCC 1460

RESULT 9

US-08-892-695-10
; Sequence 10, Application US/08892695A
; Patent No. 6808878
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W
; APPLICANT: Collins, Collin
; APPLICANT: Hwang, Soo In
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: GENES FROM THE 20Q13 AMPLICON AND THEIR USES
; FILE REFERENCE: 2500.124U53
; CURRENT APPLICATION NUMBER: US/08/892,695A
; CURRENT FILING DATE: 1997-07-15
; EARLIER APPLICATION NUMBER: 08/785,532
; EARLIER FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/731,499
; EARLIER FILING DATE: 1996-10-16
; EARLIER APPLICATION NUMBER: 08/680,395
; EARLIER FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 3186
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: ZABCI Open
; OTHER INFORMATION: Reading Frame

US-08-892-695-10

Query Match 73.6%; Score 1108.6; DB 4; Length 3186;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 1156; Conservative 3; Mismatches 28; Indels 3; Gaps 3;
Qy 320 ATGCAATCGAAGTGCAGGAACATGCAACTCAATCCCTCTTAATGTACATGAT-GG 378
Db 1 ATGCAATCGAAGTGCAGGAACATGCAACTCAATCCCTCTTAATGTACATGATGGG 60
Qy 379 CCAAGAGTGAATGGCAGCTCTCTTGGCAGTCCGATGGAGATGGA-GATGCCCTTGTCAATG 437
Db 61 CCAGAGTGAATGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATG 120
Qy 438 AAAGGCCCNCTGTTGTCAATTCGAGCTTACAAAGAAAAAATGTCAATCCGAATCGA 497

RESULT 10

Db 121 AAAGGAGCGCTGTTGTTCCATTCCGAGCTACACAAGAAAAAATGTC-ATCCAAATCGA 179
Qy 498 GGGGAATATGCCCTTTGGATTTGCATGTTCTGACGAGCAGACCTTTCACACATTTCAAGAACCT 557
Db 180 GGGGTATATGCCCTTTGGATTTGCATGTTCTGACGAGCCAGACCTTTCACACATTTCAAGAACCT 239
Qy 558 TAATAAACATGTTTAAATGCAACACCGGCTTACCCTCTGTTGAACCCAGCAGTTTCTTCGGGT 617
Db 240 TAATAAACATGTTTAAATGCAACACCGGCTTACCCTCTGTTGAACCCAGCAGTTTCTTCGGGT 299
Qy 618 TGAACGAGATATCTCAGTCCGCTTGTATTAAGTCAAGTCCGAAACAGAACCTTCCCAAGA 677
Db 300 TGAACGAGATATCTCAGTCCGCTTGTATTAAGTCAAGTCCGAAACAGAACCTTCCCAAGA 359
Qy 678 AAAGAAATTCGAAGAAAAATGAATTTAGCTGTGAGGTATGTGGGCGAGACATTTAGAGTCGC 737
Db 360 AAAGAAATTCGAAGAAAAATGAATTTAGCTGTGAGGTATGTGGGCGAGACATTTAGAGTCGC 419
Qy 738 TTTTGAATGTTGAGATCCCATGAGAACACACAAAGATTTCTTTCACTTACGGGTGTAACAT 797
Db 420 TTTTGAATGTTGAGATCCCATGAGAACACACAAAGATTTCTTTCACTTACGGGTGTAACAT 479
Qy 798 GTGCGAAGAGATTCGAAGAGCCCTTGGTTTCTTAAAAATCAATGCGGACACATATGCG 857
Db 480 GTGCGAAGAGATTCGAAGAGCCCTTGGTTTCTTAAAAATCAATGCGGACACATATGCG 539
Qy 858 CAATTCGGGGGCGCAGAACCTGACGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAA 917
Db 540 CAATTCGGGGGCGCAGAACCTGACGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAA 599
Qy 918 CGAGTCTGTCCAGGTGACGCGGCGGAGAGCATCTCTCTCTTCAAAATCTGCATGTT 977
Db 600 CGAGTCTGTCCAGGTGACGCGGCGGAGAGCATCTCTCTCTTCAAAATCTGCATGTT 659
Qy 978 TTGTGCTTCTTATTTCCAAATAAAGAAAGTCTAATTTGAGGACCCGAGGTGCACACAA 1037
Db 660 TTGTGCTTCTTATTTCCAAATAAAGAAAGTCTAATTTGAGGACCCGAGGTGCACACAA 719
Qy 1038 AAAAACTGCTTTCGGGTACGACGAGCGCGCAGACAGCATCTCCACAAGGAGGAATGCCGTC 1097
Db 720 AAAAACTGCTTTCGGGTACGACGAGCGCGCAGACAGCATCTCCACAAGGAGGAATGCCGTC 779
Qy 1098 CTCGAGGAGGACTTCTCTGCAAGTTGTTCAACTTGAGACCAAAATCTCACCTTGAACCGG 1157
Db 780 CTCGAGGAGGACTTCTCTGCAAGTTGTTCAACTTGAGACCAAAATCTCACCTTGAACCGG 839
Qy 1158 GAAGAGCTGTTCAGATGCATCCCTCAGTCCGTTCCAGTCCGTTCCAGGCTTGGCA 1217
Db 840 GAAGAGCTGTTCAGATGCATCCCTCAGTCCGTTCCAGTCCGTTCCAGGCTTGGCA 899
Qy 1218 KCTGCTACCAAGGAAGTTCCTTCCCAAGAGTGAAGGAATTTGGGCAAGAGG 1277
Db 900 GCTGCTACCAAGGAAGTTCCTTCCCAAGAGTGAAGGAATTTGGGCAAGAGG 959
Qy 1278 GAGCACCGCAACGACGATTCGAGTTCCGAGAGGAGCTTGGAGAAACAAATTAAGAACCA 1337
Db 960 GAGCACCGCAACGACGATTCGAGTTCCGAGAGGAGCTTGGAGAAACAAATTAAGAACCA 1019
Qy 1338 TTGTGAGGCTTCTCGAAGAGAAAGAGAGTGCAAAACATCTCCACGGGGAAGCGCCTC 1397
Db 1020 TTGTGAGGCTTCTCGAAGAGAAAGAGAGTGCAAAACATCTCCACGGGGAAGCGCCTC 1079
Qy 1398 CGTGACCGGATCCCAAGTTTACCCAGTAGCAAGGAGAGCCCACTCACTGCTCCGAGTG 1457
Db 1080 CGTGACCGGATCCCAAGTTTACCCAGTAGCAAGGAGAGCCCACTCACTGCTCCGAGTG 1139
Qy 1458 CGGCAAGCTTTCAAGAACCTTACCAAGTGGTCTTTCACCTCCAGGGTCC 1507
Db 1140 CGGCAAGCTTTCAAGAACCTTACCAAGTGGTCTTTCACCTCCAGGGTCC 1189

```
US-09-949-016-34699/c
; Sequence 34699, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34699
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-34699

Query Match      36.5%; Score 550.8; DB 4; Length 601;
Best Local Similarity 96.1%; Pred. No. 8.1e-173;
Matches 572; Conservative 2; Mismatches 20; Indels 1; Gaps 1;

QY 423 GATGCTTTGTCAATGAAAGGGCCNCTGTTGTCATTTCCGAGCTACACAAAGAAAAAAT 482
DB 594 GATGCTTTGTCAATGAAAGGGACCGCTGTGTTCATTTCCGAGCTACACAAAGAAAAAAT 535
QY 483 GTCATTCGAATCGAGGGGAATATGCCCTTGGAATGCAATGTTCTGCAGCCGACCTTCAC 542
DB 534 GTC-ATCCAAATCGAGGGGTATATGCCCTTGGAATGCAATGTTCTGCAGCCGACCTTCAC 476
QY 543 ACATTTCAGAACCTTTAATAAACATGTCTTAATGCAACACCGGCCCTACCTCTGTGAACC 602
DB 475 ACATTTCAGAACCTTTAATAAACATGTCTTAATGCAACACCGGCCCTACCTCTGTGAACC 416
QY 603 AGCAGTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTCGGAAC 662
DB 415 AGCAGTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTCGGAAC 356
QY 663 AGAACCTCCCAAGGAAAGAAATTCGAAGGAAATGCAAGGAAATGCAAGTATGAGGCA 722
DB 355 AGAACCTCCCAAGGAAAGAAATTCGAAGGAAATGCAAGTATGAGGCA 296
QY 723 GACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAGAACACAAAGATTTCTTTCAC 782
DB 295 GACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAGAACACAAAGATTTCTTTCAC 236
QY 783 TTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAATCAAT 842
DB 235 TTACGGGTGTAACATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAAAATCAAT 176
QY 843 CGGACRCATATATGCAAAATCGGGGCCAGAGCAAACTGCAGCAAGGCTTGGAGAGTAG 902
DB 175 CGGACACATAATGCAAAATCGGGGCCAGAGCAAACTGCAGCAAGGCTTGGAGAGTAG 116
QY 903 TCCAGCAACGATCAACAGAGGTGCTCCAGGTGCAACCGGCCGAGAGCATCTCTCTCCTTG 962
DB 115 TCCAGCAACGATCAACAGAGGTGCTCCAGGTGCAACCGGCCGAGAGCATCTCTCTCCTTA 56
QY 963 CAAAATCTGCATGTTTGTGGCTTCTCTATTTCCAAATAAAGAAAGTCTAAATTGAG 1017
DB 55 CAAAATCTGCATGTTTGTGGCTTCTCTATTTCCAAATAAAGAAAGTCTAAATTGAG 1

RESULT 11
US-09-949-016-151879/c
; Sequence 151879, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

```
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34700
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-34700

Query Match      30.2%; Score 454.8; DB 4; Length 601;
Best Local Similarity 98.3%; Pred. No. 8.9e-141;
Matches 456; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1044 TCGTTTCGTTACGAGCGCGCAGACAGACTCTCCACAGGAGGAATCCCGTCTCGAG 1103
DB 601 TCGTTTCGTTACGAGCGCGCAGACAGACTCTCCACAGGAGGAATCCCGTCTCGAG 542

QY 1104 GGAGGACTTCTCGAGTTGTTCAACTTGAGACCAAAATCTCACCCCTGAAACGGGGAAGAA 1163
DB 541 GGAGGACTTCTCGAGTTGTTCAACTTGAGACCAAAATCTCACCCCTGAAACGGGGAAGAA 482

QY 1164 GCCTGTGAGATGCATCCCTCAGCTCGATCCGTTCCAGCTTCCAGGCTTGGCAKCTGGC 1223
DB 481 GCCTGTGAGATGCATCCCTCAGCTCGATCCGTTCCAGCTTCCAGGCTTGGCAKCTGGC 422

QY 1224 TACCAAGGAAGATTGCCATTTGCCAAGAGTGAAGAAATTTGGGCAAGAGGAGGCAC 1283
DB 421 TACCAAGGAAGATTGCCATTTGCCAAGAGTGAAGAAATTTGGGCAAGAGGAGGCAC 362

QY 1284 CGACACGACGATTTCGAGTTCCGAGAGGAGCTTCGAGAAACAAATTAAGAACCATTTGTGC 1343
DB 361 CGACACGACGATTTCGAGTTCCGAGAGGAGCTTCGAGAAACAAATTAAGAACCATTTGTGC 302

QY 1344 AGCGCTCTCGCAAGAGAAAGAGAGTGCACCACTCCCAACGGCGAAGCGCCCTCCGTGGA 1403
DB 301 WGGGCTCTCGCAAGAGAAAGAGAGTGCACCACTCCCAACGGCGAAGCGCCCTCCGTGGA 242

QY 1404 CGCGATCCCAAGTTACCCAGTAGCAAGAGAGCCCACTCACTGCTCGAGTGGCGCAA 1463
DB 241 CGCGATCCCAAGTTACCCAGTAGCAAGAGAGCCCACTCACTGCTCGAGTGGCGCAA 182

QY 1464 AGCTTTTCAGAACTTACCACAGCTGCTTTGCACTCCAGGGTCC 1507
DB 181 AGCTTTTCAGAACTTACCACAGCTGCTTTGCACTCCAGGGTCC 138

RESULT 13
US-09-949-016-151880/c
; Sequence 151880, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151880
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-151880

Query Match      30.2%; Score 454.8; DB 4; Length 601;
Best Local Similarity 98.3%; Pred. No. 8.9e-141;
Matches 456; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1044 TCGTTTCGTTACGAGCGCGCAGACAGACTCTCCACAGGAGGAATCCCGTCTCGAG 1103
DB 601 TCGTTTCGTTACGAGCGCGCAGACAGACTCTCCACAGGAGGAATCCCGTCTCGAG 542

QY 1104 GGAGGACTTCTCGAGTTGTTCAACTTGAGACCAAAATCTCACCCCTGAAACGGGGAAGAA 1163
DB 541 GGAGGACTTCTCGAGTTGTTCAACTTGAGACCAAAATCTCACCCCTGAAACGGGGAAGAA 482

QY 1164 GCCTGTGAGATGCATCCCTCAGCTCGATCCGTTCCAGCTTCCAGGCTTGGCAKCTGGC 1223
DB 481 GCCTGTGAGATGCATCCCTCAGCTCGATCCGTTCCAGCTTCCAGGCTTGGCAKCTGGC 422

QY 1224 TACCAAGGAAGATTGCCATTTGCCAAGAGTGAAGAAATTTGGGCAAGAGGAGGCAC 1283
DB 421 TACCAAGGAAGATTGCCATTTGCCAAGAGTGAAGAAATTTGGGCAAGAGGAGGCAC 362

QY 1284 CGACACGACGATTTCGAGTTCCGAGAGGAGCTTCGAGAAACAAATTAAGAACCATTTGTGC 1343
DB 361 CGACACGACGATTTCGAGTTCCGAGAGGAGCTTCGAGAAACAAATTAAGAACCATTTGTGC 302

QY 1344 AGCGCTCTCGCAAGAGAAAGAGAGTGCACCACTCCCAACGGCGAAGCGCCCTCCGTGGA 1403
DB 301 WGGGCTCTCGCAAGAGAAAGAGAGTGCACCACTCCCAACGGCGAAGCGCCCTCCGTGGA 242

QY 1404 CGCGATCCCAAGTTACCCAGTAGCAAGAGAGCCCACTCACTGCTCGAGTGGCGCAA 1463
DB 241 CGCGATCCCAAGTTACCCAGTAGCAAGAGAGCCCACTCACTGCTCGAGTGGCGCAA 182

QY 1464 AGCTTTTCAGAACTTACCACAGCTGCTTTGCACTCCAGGGTCC 1507
DB 181 AGCTTTTCAGAACTTACCACAGCTGCTTTGCACTCCAGGGTCC 138

RESULT 14
US-09-513-999C-11828
; Sequence 11828, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 11828
; LENGTH: 298
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 277
; OTHER INFORMATION: s=g or c
US-09-513-999C-11828

Query Match      13.1%; Score 197.2; DB 4; Length 298;
Best Local Similarity 98.5%; Pred. No. 5e-55;
```

	Matches	199;	Conservative	0;	Mismatches	3;	Indels	0;	Gaps	0
Qy	2	CAGGTTGCTGGGAT	TGACTTCTTGCTCAAT	TGAAACACTCAT	TTCATATGAGACAAAGAGC	61				
Db	97	CAGGTTGCTGGGAT	TGACTTCTTGCTCAAT	TGAAACACTCAT	TTCATATGAGACAAAGAGAG	156				
Qy	62	ACTAATGCTTTTGCTG	TATTCATATTTGAAT	TCGAGGCAAT	TGGGAACCCCTGTATGCTTGT	121				
Db	157	ACTAATGCTTTTGCTG	TATTCATATTTGAAT	TCGAGGCAAT	TGGGAACCCCTGTATGCTTGT	216				
Qy	122	TTGTGGAAAGAA	CAAGTGACACCACT	CACCTGAGCTT	TCCTTAAAGTTTCGAAGAAAGTTAGAGG	181				
Db	217	TTGTGGAAAGAA	CAAGTGACACCACT	CACCTGAGCTT	TCCTTAAAGTTTCGAAGAAAGTTAGAGC	276				
Qy	182	ACTATACACTTTT	CTTTTGAAC	T	203					
Db	277	SCTATACACTTTT	CTTTTGAAC	T	298					

RESULT 15

```

US-09-620-312D-61
; Sequence 61, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_Flgene Version 1.0
; SEQ ID NO 61
; LENGTH: 2765
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (123)..(2291)
US-09-620-312D-61

```

	Query Match	4.6%	Score 68.8	DB 4	Length 2765	
	Best Local Similarity	62.4%	Pred. No. 1.7e-11			
	Matches	106	Conservative	1	Mismatches 63	Indels 0
					Gaps	0
Qy	697	GAATTTAGCTGTGAGGTATGTGGCCAGACATTTAGATCGCTTTTGTGATGTTGAGATCCAC	756			
Db	939	GAGTTCGCTGCCAAGTGTGGCCAGAGCTTTACACAGTCTTGTTCTCAGGGCCAC	998			
Qy	757	ATGAGAACACAAAGATTCTTTTCATTACGCGGTGTAACTATGCGCGAAGAGATTCAAG	816			
Db	999	ATGCGTTAAGCACAAAGGCTCTCTTCGATCATCGTGTCCGCTGTGCGGCCGCTGCTTCAAG	1058			

Qy 817 GAGCCTTGTTCTTAAAAATC^{AT}GCGGACR^{CA}TAA^{TG}GCAATCGG 866
||||| ||||||| ||||||| ||||||| : ||| ||| |||
Db 1059 GAGCCTTGTTCTTAAAGAA^{CC}ATGAAG^{TG}CA^{CG}CCAGCAAGCTGG 1108

Search completed: June 29, 2005, 05:27:50
Job time : 176.067 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2005, 01:06:08 ; Search time 589.497 Seconds
(without alignments)
15978.923 Million cell updates/sec

Title: US-08-731-499-3
Perfect score: 1507
Sequence: 1 GCAGTGTCTGGATTGACT.....GGTCTTGCACTCCAGGGTCC 1507

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 6067389 seqs, 3125258755 residues

Total number of hits satisfying chosen parameters: 12134778

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1:	/cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
3:	/cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
4:	/cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
7:	/cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:*
10:	/cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:*
11:	/cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
12:	/cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
13:	/cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:*
14:	/cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*
15:	/cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
16:	/cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:*
17:	/cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
18:	/cgn2_6/prodata/1/pubpna/US10F_PUBCOMB.seq:*
19:	/cgn2_6/prodata/1/pubpna/US10G_PUBCOMB.seq:*
20:	/cgn2_6/prodata/1/pubpna/US10H_PUBCOMB.seq:*
21:	/cgn2_6/prodata/1/pubpna/US10I_PUBCOMB.seq:*
22:	/cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*
23:	/cgn2_6/prodata/1/pubpna/US11A_PUBCOMB.seq:*
24:	/cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:*
25:	/cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*
26:	/cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1504.8	99.9	1507	8	US-08-731-499-3
2	1387.2	92.1	10365	8	US-08-731-499-3
3	1387.2	92.1	36022	13	US-10-087-192-1708
4	1337.2	88.7	5632	15	US-10-177-293-505
5	1132.6	75.2	5252	13	US-10-087-192-1709
6	1106.2	73.4	3186	8	US-08-731-499-10
7	573.6	38.1	26345	13	US-10-087-192-1705

RESULT 1
US-08-731-499-3
; Sequence 3, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION:
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOWBEL, David
; APPLICANT: ROMMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Townsend and Townsend and Crew
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,499
; FILING DATE: 16-OCT-1996
; CLASSIFICATION: 435

ALIGNMENTS

Sequence 1706, Ap
Sequence 9711, Ap
Sequence 520, App
Sequence 1511, Ap
Sequence 1511, Ap
Sequence 93, Appl
Sequence 23411, A
Sequence 110, App
Sequence 372, App
Sequence 20910, A
Sequence 20193, A
Sequence 61, Appl
Sequence 1644, Ap
Sequence 1396, Ap
Sequence 1394, Ap
Sequence 1396, Ap
Sequence 7003, Ap
Sequence 84, Appl
Sequence 896, App
Sequence 99, Appl
Sequence 550, App
Sequence 100, App
Sequence 9132, Ap
Sequence 225, App
Sequence 952, App
Sequence 47, Appl
Sequence 1016, Ap
Sequence 25428, A
Sequence 45, Appl
Sequence 728, App
Sequence 1916, Ap
Sequence 1628, Ap
Sequence 1628, Ap
Sequence 312, App
Sequence 1629, Ap

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/680,395
; FILING DATE: 15-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 23070-068910
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1507 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: - 1.1507
; LOCATION: 1.1507
; OTHER INFORMATION: /note= "cDNA clone cc49 of 6-7kb
; OTHER INFORMATION: transcript with homology to C2H2 zinc
; OTHER INFORMATION: finger genes"
; US-08-731-499-3

Query Match 99.9%; Score 1504.8; DB 8; Length 1507;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1507; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGGTGCTGGGATGACTCTTCTCTCAATTTGAACACTCAATCAATGGAGACAAAGAG 60
DB 1 GCAGGTGCTGGGATGACTCTTCTCTCAATTTGAACACTCAATCAATGGAGACAAAGAG 60

QY 61 CACTAATGCTTTGCTGATTCATATTTGAATTCGAGGCAATGGGAACCCCTGTATGCCCTTG 120
DB 61 CACTAATGCTTTGCTGATTCATATTTGAATTCGAGGCAATGGGAACCCCTGTATGCCCTTG 120

QY 121 TTGTGGGAAGACAGTGACACCACTACTGAGCTTCTTAAAGTTTGAAGTTAGAG 180
DB 121 TTGTGGGAAGACAGTGACACCACTACTGAGCTTCTTAAAGTTTGAAGTTAGAG 180

QY 181 GACTATACACTTCTTTTGAACCTTTTATAATAATTTGCTCTGTTTGGAAACCCAGG 240
DB 181 GACTATACACTTCTTTTGAACCTTTTATAATAATTTGCTCTGTTTGGAAACCCAGG 240

QY 241 ACTGTTAGAGGTGAGTGACAGGTCCTTACAGTGGCTTTAATCAACTCCAGAAATGGCC 300
DB 241 ACTGTTAGAGGTGAGTGACAGGTCCTTACAGTGGCTTTAATCAACTCCAGAAATGGCC 300

QY 301 AACGGAACTTTGAGATTATATGCAATCGAATGACGGAACATGCCAATCAATCCCT 360
DB 301 AACGGAACTTTGAGATTATATGCAATCGAATGACGGAACATGCCAATCAATCCCT 360

QY 361 CTTAATGTACATGATGGCCAGAGTGATGTCAGCTCTCTTCCAGTCCGATGAGATG 420
DB 361 CTTAATGTACATGATGGCCAGAGTGATGTCAGCTCTCTTCCAGTCCGATGAGATG 420

QY 421 GAGATGCTTGTCAATGAAAGGGCCNCCTGTTGTCATTTCCGAGCTACACAAAGAAAAA 480
DB 421 GAGATGCTTGTCAATGAAAGGGCCNCCTGTTGTCATTTCCGAGCTACACAAAGAAAAA 480

QY 481 ATGTCAATCCGATCGAGGGGAATATGCCCTGGATTGCAATGCTTGCAGCCAGACCTTC 540
DB 481 ATGTCAATCCGATCGAGGGGAATATGCCCTGGATTGCAATGCTTGCAGCCAGACCTTC 540

QY 541 ACACATTGAGAGACCTTTAATAAATGCTTTAATGCAACACCGGCTTACCCTCTGTGAA 600
DB 541 ACACATTGAGAGACCTTTAATAAATGCTTTAATGCAACACCGGCTTACCCTCTGTGAA 600

QY 601 CCAGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGTATATAAGTCAAGTGGGA 660
DB 601 CCAGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGTATATAAGTCAAGTGGGA 660

```

```

QY 661 ACAGAACCTCCCAAGGAAAGAAATTTGCAAGAAATGAATTTAGCTGTGAGGTATGTGGG 720
DB 661 ACAGAACCTCCCAAGGAAAGAAATTTGCAAGAAATGAATTTAGCTGTGAGGTATGTGGG 720

QY 721 CAGACATTTAGAGTGCCTTTTGTGATGTTGAGATCCACATGAGAACACACAAAGATTTCTTC 780
DB 721 CAGACATTTAGAGTGCCTTTTGTGATGTTGAGATCCACATGAGAACACACAAAGATTTCTTC 780

QY 781 ACTTACGGGTGTAAATGCGGAAAGAGATTTCAAGGAGCCTTGTGTTCTTAAATATCAC 840
DB 781 ACTTACGGGTGTAAATGCGGAAAGAGATTTCAAGGAGCCTTGTGTTCTTAAATATCAC 840

QY 841 ATGGGACRCATATATGCGCAATTCGGGGCCAGAGCAAACTGCAGCAAGGCTTGAGAGT 900
DB 841 ATGGGACRCATATATGCGCAATTCGGGGCCAGAGCAAACTGCAGCAAGGCTTGAGAGT 900

QY 901 AGTCCAGCAACGATCAACGAGGTCTGTCAGGTGCAACGCGGCGGAGAGCATCTCTCTCT 960
DB 901 AGTCCAGCAACGATCAACGAGGTCTGTCAGGTGCAACGCGGCGGAGAGCATCTCTCTCTCT 960

QY 961 TCGAAATCTGATGTTGTTGCTTCTTCTTCAATTAAGAAAGTCTAATTTAGACAC 1020
DB 961 TCGAAATCTGATGTTGTTGCTTCTTCTTCAATTAAGAAAGTCTAATTTAGACAC 1020

QY 1021 CGCAGGTGCAACCAAAATGCTTTTCTGTCAGCGGCGGAGAGCATCTCTCTCTCTCA 1080
DB 1021 CGCAGGTGCAACCAAAATGCTTTTCTGTCAGCGGCGGAGAGCATCTCTCTCTCTCA 1080

QY 1081 CAAGGAGGAATGCGCTCTCTCGAGGAGGAGCTTCTCTGAGTTGTTTCAACTTGAGACCAAAA 1140
DB 1081 CAAGGAGGAATGCGCTCTCTCGAGGAGGAGCTTCTCTGAGTTGTTTCAACTTGAGACCAAAA 1140

QY 1141 TCTCACTCTGAAAACGGGAAAGAGCTGTGATGATCCTCTCAGCTCGATCCGTTTCAAC 1200
DB 1141 TCTCACTCTGAAAACGGGAAAGAGCTGTGATGATCCTCTCAGCTCGATCCGTTTCAAC 1200

QY 1201 ACCTTCCAGGCTTGCACTGCTTACCAAGAAAGTTGCCATTTGCCAAGAGTGAAG 1260
DB 1201 ACCTTCCAGGCTTGCACTGCTTACCAAGAAAGTTGCCATTTGCCAAGAGTGAAG 1260

QY 1261 GAAATGGGGCAAGAGGAGGAGCAACGAAACGATTCGAGTTCCGAGAAGAGCTTGA 1320
DB 1261 GAAATGGGGCAAGAGGAGGAGCAACGAAACGATTCGAGTTCCGAGAAGAGCTTGA 1320

QY 1321 GAAACAAATAGAACCTTGTGAGGCTCTCGAAGAAAGAGAGTGCACAACTCC 1380
DB 1321 GAAACAAATAGAACCTTGTGAGGCTCTCGAAGAAAGAGAGTGCACAACTCC 1380

QY 1381 CACGGCAAGCGCCCTCCGTCGACGCGATCCCAAGTTTACCCAGTAGCAAGAGAGGCC 1440
DB 1381 CACGGCAAGCGCCCTCCGTCGACGCGATCCCAAGTTTACCCAGTAGCAAGAGAGGCC 1440

QY 1441 ACTCACTGCTCCGAGTGGCAAAAGCTTTTCAGAACTTACCACCTAGCTGCTTGCATCTCC 1500
DB 1441 ACTCACTGCTCCGAGTGGCAAAAGCTTTTCAGAACTTACCACCTAGCTGCTTGCATCTCC 1500

QY 1501 AGGTTCC 1507
DB 1501 AGGTTCC 1507

```

```

RESULT 2
US-08-731-499-9
; Sequence 9, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION:
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOWBEL, David
; APPLICANT: ROMMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR

```

1
TITLE OF INVENTION: USES
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/731,499
FILING DATE: 16-OCT-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/680,395
FILING DATE: 15-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 23070-068910
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 10365 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: -
LOCATION: 1..10365
OTHER INFORMATION: /note= "Genomic Sequence Encoding
OTHER INFORMATION: ZABC1"
US-08-731-499-9

Query Match 92.1%; Score 1387.2; DB 8; Length 10365;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1470; Conservative 3; Mismatches 32; Indels 6; Gaps 6;
QY 2 CAGGTTCCTGGGATTCGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAGAGC 61
DB 2424 CAGGTTCCTGGGATTCGACTTCTTGCTCAATTGAAACACTCATTCAATGGAGACAAGAGA 2483
QY 62 ACTAATGCTTTTGCTGATTCATATTGAAATCGAGGCATTGGGAACCCCTGTATGCTTGT 121
DB 2484 ACTAATGCTTTTGCTGATTCATATTGAAATCGAGGCATTGGGAACCCCTGTATGCTTGT 2543
QY 122 TTGTGAAAGAACCGATGACACATCCTGAGCTTCTTAAAGATTCGAAAGAGTTAGAGG 181
DB 2544 TTGTGAAAGAACCGATGACACATCCTGAGCTTCTTAAAGATTCGAAAGAGTTAGAGG 2603
QY 182 ACTATACACTTCTTTTGAACTTTTATATAATATTGCTCTGG-TTTTGGAAACCCAGG 240
DB 2604 ACTATACACTTCTTTTGAACTTTTATATAATATTGCTCTGGTTTGGAAACCCAGG 2663
QY 241 ACTGTTTGA-GGGTGAGTGACAGGCTTAC-AGTGGCCCTTAATCCAACCTCCAGAAATTGC 298
DB 2664 GCTGTTAGGGGGTGGTGACAGCTTACAGTGGCCCTTATCCAACTCCAGAAATTGC 2723
QY 299 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 358
DB 2724 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCC 2783
QY 359 CTCCTTAATGTACATGGAT-GGGCAAGAGTGATGGCAGCTCTCTTGCCAGTCCGATGGAG 417
DB 2784 CTCCTTAATGTACATGGATGGGCCAGAGTGATGGCAGCTCTCTTGCCAGTCCGATGGAG 2843

QY 418 ATGGA-GATGCCTTGTCAATGAAAGGGCCNCTGTTGTCAATTCGAGCTACACAAAGAA 476
DB 2844 ATGGAGGATGCTTGTCAATGAAAGGGACCGCTGTTGTTCCATTCGAGCTACACAAAGAA 2903
QY 477 AAAAATGTCATCCGAATCGAGGGGAATATGCCCTTGGATTGATGATGTTCTGAGCCAGAC 536
DB 2904 AAAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTGGATTGATGATGTTCTGAGCCAGAC 2962
QY 537 CTTTCACACATTCAGAAAGACCTTAATAAATCATGCTTAATGCAACACCGCCCTACCCCTCTG 596
DB 2963 CTTTCACACATTCAGAAAGACCTTAATAAATCATGCTTAATGCAACACCGCCCTACCCCTCTG 3022
QY 597 TGAACACAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGAATAAAGTCAAGT 656
DB 3023 TGAACACAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGAATAAAGTCAAGT 3082
QY 657 GCGAACAGACCTCCNAGGAAAGAAATGCAAGGAAATGAATTTAGCTGTGAGGTATG 716
DB 3083 GCGAACAGACCTCCNAGGAAAGAAATGCAAGGAAATGAATTTAGCTGTGAGGTATG 3142
QY 717 TGGGCAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAGAAACACACAAAGATTTC 776
DB 3143 TGGGCAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAGAAACACACAAAGATTTC 3202
QY 777 TTTCACTTACGGGTGTAAATGTCGGGAAGAAAGATTCAAGGAGCCTTGGTTCTTTAAAAA 836
DB 3203 TTTCACTTACGGGTGTAAATGTCGGGAAGAAAGATTCAAGGAGCCTTGGTTCTTTAAAAA 3262
QY 837 TCACATGCGGACRCAATATGCGCAATCGGGGCCAGAGCAAACTGCAGCAGGCTTGA 896
DB 3263 TCACATGCGGACACATAATGCGCAATCGGGGCCAGAGCAAACTGCAGCAGGCTTGA 3322
QY 897 GAGTAGTCAGCAACGATCAACGAGTGTCTCAGAGTGCAACGCGCGCGAGAGCATCTCCTC 956
DB 3323 GAGTAGTCAGCAACGATCAACGAGTGTCTCAGAGTGCAACGCGCGCGAGAGCATCTCCTC 3382
QY 957 TCCTTGCAAAATCTGCATGTTGTTGGTTCCTATTTCAAAATAAGAAAGTCTAATTGA 1016
DB 3383 TCCTTGCAAAATCTGCATGTTGTTGGTTCCTATTTCAAAATAAGAAAGTCTAATTGA 3442
QY 1017 GCACGCGAGGTGACACCAAAAACCTGCTTCGTTACAGCAGCGCGCGAGAGCATCTC 1076
DB 3443 GCACGCGAGGTGACACCAAAAACCTGCTTCGTTACAGCAGCGCGCGAGAGCATCTC 3502
QY 1077 TCACAGAGGAAATGCGCTCTCGAGGAGAGCTTCTCTGCAAGTGTGTTCAACTTGAGACC 1136
DB 3503 TCACAGAGGAAATGCGCTCTCGAGGAGAGCTTCTCTGCAAGTGTGTTCAACTTGAGACC 3562
QY 1137 AAAATCTCACCTGAAAACGGGGAAGAGCCTGTGAGATGCAATCCCTCAGCTCGATCCGTT 1196
DB 3563 AAAATCTCACCTGAAAACGGGGAAGAGCCTGTGAGATGCAATCCCTCAGCTCGATCCGTT 3622
QY 1197 CACCACTTCCAGGCTTGGCAKCTGCTTACCAAGAAAGAGTTCATTTGCAAGAGT 1256
DB 3623 CACCACTTCCAGGCTTGGCAKCTGCTTACCAAGAAAGAGTTCATTTGCAAGAGT 3682
QY 1257 GAAGCAATTTGGGGAAGAGGAGCAGCAGCAACGATTCGAGTTCGGAAGAGAGCT 1316
DB 3683 GAAGCAATTTGGGGAAGAGGAGCAGCAGCAGCAACGATTCGAGTTCGGAAGAGAGCT 3742
QY 1317 TGGAGAAACAAATAGAACCAATTTGTGAGGCTCTCGCAAGAGAAAGAGAGTGCAGAA 1376
DB 3743 TGGAGAAACAAATAGAACCAATTTGTGAGGCTCTCGCAAGAGAAAGAGAGTGCAGAA 3802
QY 1377 CTCACGCGAGAGCGCCCTCGTGGAACGCGATCCCAAGTTACCAGTAGCAAGAGAA 1436
DB 3803 CTCACGCGAGAGCGCCCTCGTGGAACGCGATCCCAAGTTACCAGTAGCAAGAGAA 3862
QY 1437 GCCCACTCACTGCTCCGAGTGGGCAAGAGCTTTAGAACCTTACCAACCTAGCTGCTTGA 1496
DB 3863 GCCCACTCACTGCTCCGAGTGGGCAAGAGCTTTAGAACCTTACCAACCTAGCTGCTTGA 3922

```
QY 1497 CTCACGGGTCC 1507
|||||
Db 3923 CTCACGGGTCC 3933

RESULT 3
US-10-087-192-1708
; Sequence 1708, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE OF INVENTION: CANCER
; FILE REFERENCE: 52945200122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1708
; LENGTH: 36022
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1708

Query Match 92.1%; Score 1387.2; DB 13; Length 36022;
Best Local Similarity 97.3%; Pred. No. 0;
Matches 1470; Conservative 3; Mismatches 32; Indels 6; Gaps 6;

QY 2 CAGGTTCGTGGATTGACTTCTTGCTCAATTGAAACACTCAATTCAATGGAGCAAAAGAGC 61
Db 9951 CAGGTTCGTGGATTGACTTCTTGCTCAATTGAAACACTCAATTCAATGGAGCAAAAGAGA 10010

QY 62 ACTAATGCTTTGTGCTGATTTCATATTTGAATCGAGGCATTCGGACCCCTGTATGCTTGT 121
Db 10011 ACTAATGCTTTGTGCTGATTTCATATTTGAATCGAGGCATTCGGACCCCTGTATGCTTGT 10070

QY 122 TTGTGAAAGAACCAAGTGACACCACTCACTGAGCTTCTTAAAGTTTCGAAGAGTTAGAGG 181
Db 10071 TTGTGAAAGAACCAAGTGACACCACTCACTGAGCTTCTTAAAGTTTCGAAGAGTTAGAGG 10130

QY 182 ACTATACACTTTCTTTTGAACTTTTATATAAATTTTGTCTCTG- TTTTGAACCCAGG 240
Db 10131 ACTATACACTTTCTTTTGAACTTTTATATAAATTTTGTCTCTGTTTGTGAACCCAGG 10190

QY 241 ACTGTTTGA- GGGTGAGTGACAGGTCTTAC- AGTGGCCCTTAATCCAACTCCAGAAATTC 298
Db 10191 GCTGTTAGAGGGTGAGTGACAAAGTCTTACAAAGTTCAGAAAGTTCAGAAATTC 10250

QY 299 CCAACGGAACCTTGAGATTATATGCAATCGAAAGTGACAGAAACATGCAACTCAATCC 358
Db 10251 CCAACGGAACCTTGAGATTATATGCAATCGAAAGTGACAGAAACATGCAACTCAATCC 10310

QY 359 CTCCTTAATGTACATGGAT- GGCACAGAGTGAATGGCAGCTCTCTTGGCCAGTCCGATGAG 417
Db 10311 CTCCTTAATGTACATGGATGGGCGCAGAAAGTGAATGGCAGCTCTCTTGGCCAGTCCGATGAG 10370

QY 418 ATGGA- GATGCTTCTCAATGAAGGGCCCTGTTGTCAATTCGAGCTACACAAAGA 476
Db 10371 ATGGAAGATGCTTCTCAATGAAGGGCCCTGTTGTTCATTCGAGCTACACAAAGA 10430

QY 477 AAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCAATGTTCTGAGCCAGAC 536
Db 10431 AAAAATGTC- ATCCAAATCGAGGGGTATATGCCCTTGGATTGCAATGTTCTGAGCCAGAC 10489

QY 537 CTTGACACATTCAGAAAGACCTTAATAAATGCTCTTAATGCAACACCGGCTTACCTCTG 596
Db 10490 CTTGACACATTCAGAAAGACCTTAATAAATGCTCTTAATGCAACACCGGCTTACCTCTG 10549
```

```
QY 597 TGAACACGACAGTCTCTTGGGTTGAAGCAGAGTATCTCAGTCGGCTTGATATAAGTCAAGT 656
|||||
Db 10550 TGAACACGACAGTCTCTTGGGTTGAAGCAGAGTATCTCAGTCGGCTTGATATAAGTCAAGT 10609

QY 657 GCGAACAGAACCTCCCAAGGAAAAGAAATGCAAGGAAAATGAATTTAGCTGTGAGGTATG 716
|||||
Db 10610 GCGAACAGAACCTCCCAAGGAAAAGAAATGCAAGGAAAATGAATTTAGCTGTGAGGTATG 10669

QY 717 TGGGCGAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAGAACACACAAAGATTC 776
|||||
Db 10670 TGGGCGAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATGAGAACACACAAAGATTC 10729

QY 777 TTTCACTTTACGGGTGAACATGTGCGGAAGAAAGATTCAAGGAGCCCTGCTTTCTTAAAAA 836
|||||
Db 10730 TTTCACTTTACGGGTGAACATGTGCGGAAGAAAGATTCAAGGAGCCCTGCTTTCTTAAAAA 10789

QY 837 TCACATGCGGACRCATATATGGCAAAATCGGGGGCCAGAAAGCAAACTGACAGAGGCTTGA 896
|||||
Db 10790 TCACATGCGGACACATATATGCAAAATCGGGGGCCAGAAAGCAAACTGACAGAGGCTTGA 10849

QY 897 GAGTAGTCAGAACGATCAACGAGGTCTGTCAGGTGACCGGGCCGAGAGCATCTCTC 956
|||||
Db 10850 GAGTAGTCAGAACGATCAACGAGGTCTGTCAGGTGACCGGGCCGAGAGCATCTCTC 10909

QY 957 TCCTTGCAAAATCTGCATGTTTGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTGA 1016
|||||
Db 10910 TCCTTGCAAAATCTGCATGTTTGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTGA 10969

QY 1017 GCACCGCAAGGTGCACACCAAAAAAACTGCTTTTGGTACCAGCAGCGCGCAGACAGACTC 1076
|||||
Db 10970 GCACCGCAAGGTGCACACCAAAAAAACTGCTTTTGGTACCAGCAGCGCGCAGACAGACTC 11029

QY 1077 TCCACAGGAGGAATGCGCTCTCGAGGAGGACTTCTCTGACATGTTGTTCAACTTGAGACC 1136
|||||
Db 11030 TCCACAGGAGGAATGCGCTCTCGAGGAGGACTTCTCTGACATGTTGTTCAACTTGAGACC 11089

QY 1137 AAAATCTCACCTCGAAACGGGGAGAAAGCCCTGTGATGTCATCCCTCAGCTCGATCCGTT 1196
|||||
Db 11090 AAAATCTCACCTGAAACGGGGAGAAAGCCCTGTGATGTCATCCCTCAGCTCGATCCGTT 11149

QY 1197 CACCACTTCCAGGCTTGGCAKCTGGCTACAAAGGAAAGTTCATTTGCCAAGAAAGT 1256
|||||
Db 11150 CACCACTTCCAGGCTTGGCAGCTGGCTACCAAGGAAAGTTCATTTGCCAAGAAAGT 11209

QY 1257 GAAAGATTTGGGGCAAGAGGAGCAGCCGACAAACGATTCGAGTTCGAGAGGAGCT 1316
|||||
Db 11210 GAAAGATTTGGGGCAAGAGGAGCAGCCGACAAACGATTCGAGTTCGAGAGGAGCT 11269

QY 1317 TGGAGAAACAAATAAGAACCACTTGTGAGGCTCTCGCAAGAGAAAGAGAGTGCAAAACA 1376
|||||
Db 11270 TGGAGAAACAAATAAGAGGCACTTGTGAGGCTCTCGCAAGAGAAAGAGAGTGCAAAACA 11329

QY 1377 CTCCTCACGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAGCAAGAGAA 1436
|||||
Db 11330 CTCCTCACGCGAAGCGCCCTCCGTGGACGCGGATCCCAAGTTACCCAGTAGCAAGAGAA 11389

QY 1437 GCCCACTCAGTCTCGAGTCGCGGCAAGCTTTCAGAACCTTACCACAGCTGGTCTTGA 1496
|||||
Db 11390 GCCCACTCAGTCTCGAGTCGCGGCAAGCTTTCAGAACCTTACCACAGCTGGTCTTGA 11449

QY 1497 CTCACGGGTCC 1507
|||||
Db 11450 CTCACGGGTCC 11460
```

```
RESULT 4
US-10-177-293-505
; Sequence 505, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
```

APPLICANT: Gannavarpu, Manjula
APPLICANT: Kamatkar, Shaubhangi
APPLICANT: Mertens, Maureen
APPLICANT: Myer, Vic
APPLICANT: Wang, Youzhen
APPLICANT: Xu, Yongyao
APPLICANT: Hoersch, Sebastian
APPLICANT: Monahan, John
APPLICANT: Meyers, Rachel E.
APPLICANT: Bast Jr., Robert C.
APPLICANT: Hortobagyi, Gabriel N.
APPLICANT: Pusztai, Lajos
APPLICANT: Meric, Funda
APPLICANT: Sahin, Aysegul
APPLICANT: Mills, Gordon B.
TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
PREVENTION, AND THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-038
CURRENT APPLICATION NUMBER: US/10/177,293
CURRENT FILING DATE: 2002-06-21
PRIOR APPLICATION NUMBER: US 60/299,887
PRIOR FILING DATE: 2001-06-21
PRIOR APPLICATION NUMBER: US 60/301,572
PRIOR FILING DATE: 2001-06-27
PRIOR APPLICATION NUMBER: US 60/306,501
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: US 60/325,002
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/362,585
PRIOR FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: US 60/xxx,xxx
PRIOR FILING DATE: 2002-05-14
NUMBER OF SEQ ID NOS: 506
SOFTWARE: PastSeq for Windows Version 4.0
SEQ ID NO 505
LENGTH: 5632
TYPE: DNA
ORGANISM: Homo sapiens
US-10-177-293-505

Query Match 88.7%; Score 1337.2; DB 15; Length 5632;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1420; Conservative 3; Mismatches 32; Indels 6; Gaps 6;
QY 52 GACAAAGACCTAATGCTTTGGCTGATTCATATTTGAATCGAGCATTTGGGACCCCTG 111
DB 1 GACAAAGACCTAATGCTTTGGCTGATTCATATTTGAATCGAGCATTTGGGACCCCTG 60
QY 112 TATGCTTTGTTGGAAAGAACCAAGTGCACCATCCTGAGCTTCTTAAAGTTGGAAG 171
DB 61 TATGCTTTGTTGGAAAGAACCAAGTGCACCATCCTGAGCTTCTTAAAGTTGGAAG 120
QY 172 AGTTAGAGCATATACATTTCTTTTGAATTTTAACTTTTAAATTAATTTGCTCTGG-TTTT 230
DB 121 AGTTAGAGCATATACATTTCTTTTGAATTTTAAATTAATTTGCTCTGGTTTTT 180
QY 231 GGAACCCAGACTGTTTGA-GGTGAGTGCACAGGCTTAC-AGTGCCCTTAATCCAACTC 288
DB 181 GGAACCCAGGCTTTAGAGGGGTGAGTGCACAGTCTTACAGTGGCTTTATCCAACTC 240
QY 289 CAGAAATGCCCAACGGAATTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 348
DB 241 CAGAAATGCCCAACGGAATTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCC 300
QY 349 AACTCAATCCCTTTAATGTACATGAT-GGGCAGAGTGTGGCAGCTCTTGGCCAG 407
DB 301 AACTCAATCCCTTTAATGTACATGATGGGCGAGAGTGTGGCAGCTCTTGGCCAG 360
QY 408 TCCGATGGAGATGGA-GATGCCCTTTGCAATGAAGGCCCNCTGTTGTCAATTCGAGCT 466
DB 361 TCCGATGGAGATGGAGATGGCTTTGTCAATGAAGGACCGCTGTTGTTCCATTCGAGC 420
QY 467 ACACAAAGAAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCT 526

DB 421 TACACAAGAAAAAATGTC-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCT 479
QY 527 GCAGCCAGACCTTTACACATTCCAGAACCTTTAAATAAACATGCTCTTAATGCAACACCGGC 586
DB 480 GCAGCCAGACCTTTACACATTCCAGAACCTTTAAATAAACATGCTCTTAATGCAACACCGGC 539
QY 587 CTACCCCTCTGTGAACCCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCGCTTTGATA 646
DB 540 CTACCCCTCTGTGAACCCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCGCTTTGATA 599
QY 647 AAAGTCAAGTGCAGAACAGAACCTCCCAAGGAAAAAGATTGCAAGGAAAAATGAATTTAGCT 706
DB 600 AAAGTCAAGTGCAGAACAGAACCTCCCAAGGAAAAAGATTGCAAGGAAAAATGAATTTAGCT 659
QY 707 GTGAGGTATGTGGGAGACATTTAGAGTCGCTTTTGAATGTTGAGATCCACATGAGAACAC 766
DB 660 GTGAGGTATGTGGGAGACATTTAGAGTCGCTTTTGAATGTTGAGATCCACATGAGAACAC 719
QY 767 ACAAAGATTCTTTCACTTACCGGTTGTAAACATGTGGGGAAGAGATTCAAGGAGCCTTGGT 826
DB 720 ACAAAGATTCTTTCACTTACCGGTTGTAAACATGTGGGGAAGAGATTCAAGGAGCCTTGGT 779
QY 827 TTCTTAAAAATCACATGGGACRCATATATGGCAAAATCGGGGGCCAGAACCAAACTGCAGC 886
DB 780 TTCTTAAAAATCACATGGGACRCATATATGGCAAAATCGGGGGCCAGAACCAAACTGCAGC 839
QY 887 AAGGCTTGGAGAGTAGTCCAGCAACGATCAACAGAGTCGTCCAGGTGACGCGGCGCGAGA 946
DB 840 AAGGCTTGGAGAGTAGTCCAGCAACGATCAACAGAGTCGTCCAGGTGACGCGGCGCGAGA 899
QY 947 GCATCTCTCTCTTTCACAAAATCTGCATGTTTGGCTTCTCTATTTCACAAATAAAGAAA 1006
DB 900 GCATCTCTCTCTTTCACAAAATCTGCATGTTTGGCTTCTCTATTTCACAAATAAAGAAA 959
QY 1007 GTCTAATTGAGCACCGCAAGGTGCAACCAAAAAAATCTGCTTTTGGTACGAGGCGCGC 1066
DB 960 GTCTAATTGAGCACCGCAAGGTGCAACCAAAAAAATCTGCTTTTGGTACGAGGCGCGC 1019
QY 1067 AGACAGACTCTCCACAAGGAGGAATGCGCTCTCGAGGGAGGACTTCTCGCAGTTGTTCA 1126
DB 1020 AGACAGACTCTCCACAAGGAGGAATGCGCTCTCGAGGGAGGACTTCTCGCAGTTGTTCA 1079
QY 1127 ACTTGAGACCAAAATCTCACCCCTGAAACGGGGAAGAGCCTGTGAGATGCATCCCTCAGC 1186
DB 1080 ACTTGAGACCAAAATCTCACCCCTGAAACGGGGAAGAGCCTGTGAGATGCATCCCTCAGC 1139
QY 1187 TCGATCCGTTTACCACTTTCAGGCTTGGCAKCTGGCTACCAAGAAWAGTTGCCATTT 1246
DB 1140 TCGATCCGTTTACCACTTTCAGGCTTGGCAKCTGGCTACCAAGAAWAGTTGCCATTT 1199
QY 1247 GCCAAGAGTGAAGGAATTTGGGCAAGAGGAGGAGCAGCAACGACGATTCGAGTTCCG 1306
DB 1200 GCCAAGAGTGAAGGAATTTGGGCAAGAGGAGGAGCAGCAACGACGATTCGAGTTCCG 1259
QY 1307 AGAAGGAGCTTGGAGAAAACAAATAAGAACCATTTGTGCGAGGCTCTCGCAAGAGAAAGAGA 1366
DB 1260 AGAAGGAGCTTGGAGAAAACAAATAAGAACCATTTGTGCGAGGCTCTCGCAAGAGAAAGAGA 1319
QY 1367 AGTGCAAAACATCTCCACCGGCAAGGCGCTTCGTTGGAACGCGGATCCCAAGTTACCCAGTA 1426
DB 1320 AGTGCAAAACATCTCCACCGGCAAGGCGCTTCGTTGGAACGCGGATCCCAAGTTACCCAGTA 1379
QY 1427 GCAAGGAGAGCCCACTCACTGCTCCGAGTGGGCAAGAGCTTTTCAGAACTTACCACCGAC 1486
DB 1380 GCAAGGAGAGCCCACTCACTGCTCCGAGTGGGCAAGAGCTTTTCAGAACTTACCACCGAC 1439
QY 1487 TGGTCTTGCATCTCCAGGGTCC 1507
DB 1440 TGGTCTTGCATCTCCAGGGTCC 1460

RESULT 5

```

US-10-087-192-1709
; Sequence 1709, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1709
; LENGTH: 5252
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1709

Query Match      75.2%; Score 1132.6; DB 13; Length 5252;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 1180; Conservative 3; Mismatches 28; Indels 3; Gaps 3;

QY 296 TGCCCAACGGAACTTTGAGATTATATGCAATGGAAGTGCAGAGAAACATGCCAACTCAA 355
DB 1 TGCCCAACGGAACTTTGAGATTATATGCAATGGAAGTGCAGAGAAACATGCCAACTCAA 60

QY 356 TCCCTCTTAATGTACATCGAT-GGCCAAGAGTGTGGCAGCTCTCTTGGCCAGTCCGATG 414
DB 61 TCCCTCTTAATGTACATCGATGGCCAGAGTGTGGCAGCTCTCTTGGCCAGTCCGATG 120

QY 415 GAGATGGA-GATGCCCTTGTCAATGAAGGGCCNCTGTTGTCAATTCGAGCTACACAA 473
DB 121 GAGATGGAAGTGCCTTGTCAATGAAGGGCCGCTGTTGTCAATTCGAGCTACACAA 180

QY 474 GAAAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCTGAGCCA 533
DB 181 GAAAAAATGTG-C-ATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGAGCCA 239

QY 534 GACCTTCACATTCAGAGACCTTAATAACATGCTTAATGCAACACCGCCCTACCCCT 593
DB 240 GACCTTCACATTCAGAGACCTTAATAACATGCTTAATGCAACACCGCCCTACCCCT 299

QY 594 CTGTGAACCCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCA 653
DB 300 CTGTGAACCCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCA 359

QY 654 AGTGGGAACAGAACCTCCAGGAAAGAAATGCAAGAAATGAATTTAGTGTGAGGT 713
DB 360 AGTGGGAACAGAACCTCCAGGAAAGAAATGCAAGAAATGAATTTAGTGTGAGGT 419

QY 714 ATGTGGGAGACATTTAGAGTCGCTTTTGTGATGTTTCAGATCCACATGAGAACACAAAGA 773
DB 420 ATGTGGGAGACATTTAGAGTCGCTTTTGTGATGTTTCAGATCCACATGAGAACACAAAGA 479

QY 774 TTCTTTTCACTTACCGGTTGAATGCGGGAAGAAATTCAGGAGCCCTTGGTTTCTTAA 833
DB 480 TTCTTTTCACTTACCGGTTGAATGCGGGAAGAAATTCAGGAGCCCTTGGTTTCTTAA 539

QY 834 AAATCAGATGGGACRCATATATGGAATCGGGGCCAGAGCAAACTGCAGCAGGCTT 893
DB 540 AAATCAGATGGGACRCATATATGGAATCGGGGCCAGAGCAAACTGCAGCAGGCTT 599

QY 894 GGAGAGTAGTCAGCAACGATCAAGAGTCTCTCAGGTGCACGGCCGAGAGCATCTC 953
DB 600 GGAGAGTAGTCAGCAACGATCAAGAGTCTCTCAGGTGCACGGCCGAGAGCATCTC 659

QY 954 CTCTCTTTGCAAAATCTCGATGGTTTGTGGCTTCTATTTTCCAAATAAAGAAAGTCTAAT 1013
DB 954 CTCTCTTTGCAAAATCTCGATGGTTTGTGGCTTCTATTTTCCAAATAAAGAAAGTCTAAT 1013

```

```

DB 660 CTCTCTTTACAAAATCTGCATGTGTTTGTGGCTTCTCTATTTCCAAAATAAAGAAAGTCTAAT 719
QY 1014 TGAGCAGCGAAGTGTGCACACCAAAAACCTGCTTTGGTACCAGCAGCGCGCAGACAGA 1073
DB 720 TGAGCAGCGAAGTGTGCACACCAAAAACCTGCTTTGGTACCAGCAGCGCGCAGACAGA 779
QY 1074 CTCTCCCAAGGAGGAATGCGCTTCTCGAGGAGGACTTCTGCAAGTTGTTCAACTTCGAG 1133
DB 780 CTCTCCCAAGGAGGAATGCGCTTCTCGAGGAGGACTTCTGCAAGTTGTTCAACTTCGAG 839
QY 1134 ACCAAAATCTCACCTGAAACGGGGAAGAACCTGTGCAGATGCATCCCTCAGCTCGATCC 1193
DB 840 ACCAAAATCTCACCTGAAACGGGGAAGAACCTGTGCAGATGCATCCCTCAGCTCGATCC 899
QY 1194 GTTACCAACCTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGAAAGTTGCCATTTGCCAAGA 1253
DB 900 GTTACCAACCTTCCAGGCTTGGCAKCTGGCTACCAAGGAAGAAAGTTGCCATTTGCCAAGA 959
QY 1254 AGTGAAGGAATTTGGGGCAAGGAGCACCACAAACGACGATTCGAGTTCGGAGAAGA 1313
DB 960 AGTGAAGGAATTCGGGGCAAGGAGCACCACAAACGACGATTCGAGTTCGGAGAAGA 1019
QY 1314 GCTTGGAGAAAACAAATAAGAACCTTGTGCAAGGCTTCTGCAAGGAAGAAAGTGCAG 1373
DB 1020 GCTTGGAGAAAACAAATAAGGCGAGTTGTGCAAGGCTTCTGCAAGGAAGAAAGTGCAG 1079
QY 1374 ACATCTCCACGGGAGCGCCCTCGGTGAGCGCGATCCCAAGTTACCCAGCTAGCAAGA 1433
DB 1080 ACATCTCCACGGGAGCGCCCTCGGTGAGCGCGATCCCAAGTTACCCAGCTAGCAAGA 1139
QY 1434 GAAGCCCACTCACTGCTCCGAGTGGCGCAAAAGCTTTCAAGAACCTTACCAACCTGCTT 1493
DB 1140 GAAGCCCACTCACTGCTCCGAGTGGCGCAAAAGCTTTCAAGAACCTTACCAACCTGCTT 1199
QY 1494 GCATCCAGGTGCC 1507
DB 1200 GCATCCAGGTGCC 1213

RESULT 6
US-08-731-499-10
; Sequence 10, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION:
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOWBEL, David
; APPLICANT: ROMMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR
; TITLE OF INVENTION: USES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,499
; FILING DATE: 16-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: US 08/680,395
; FILING DATE: 15-JUL-1996
; ATTORNEY/AGENT INFORMATION:

```


Db 9839 AATATACTCGTCCCTCTGACCTT-----AGAGGTGCTCTGATTTTGACACCCAGG 9889
 Qy 242 CTGTTAGAGGTGAG---TGACAGTCTTACAGTGGCCTTAATCAACTCCAGAAATTCG 298
 Db 9890 CTCCTTAGAGGTCTGAGTGGTGATATCTCACAGGCGGCTTATTCTAACTGCAGAGATCC 9949
 Qy 299 CCAACGGAACTTTGAGATTATATGCAATCGA-----AAGTGACAGGAAACATGCCAACT 352
 Db 9950 TCAGCAGATAGTGGATTACCCATCTCTGACACCTCTGAAGTGGTGGCAGCATGCCGACC 10009
 Qy 353 CAATCCCTTTAATGATACATGA-TGGCCAAAGATGATTGGCAGCTCTCTTGGCCAGTCCG 411
 Db 10010 CAGTCCCTCTCGTGATCATGACGCGCGGGAAGTCTCTCAGCAGCTCTCTAGGCTCCAG 10069
 Qy 412 ATGGAGATGGAGATGCCCTTGTCAATGAAGGCCCTCTGTTGTCAATTCGAGCTACACA 471
 Db 10070 ATGGAGGTGGATGATGCTGTGCCCCATGAAGGCCGCTGGCAGTCCCCCTTCGAGCTGCT 10129
 Qy 472 AAGAAAAAATGTCATCCGAATCGAGGGGAATATGCCCTTGGATTGCAATGCTTCGACGC 531
 Db 10130 CAGGAGAGCATGGCCGTGGCAGAGGGCCACATGCCCTGGATTGCAATGCTTCGACGC 10189
 Qy 532 CAGACCTTCACACATTCAGAAAGACCTTAATAACATGCTTTAATGAACACCGGCTTACC 591
 Db 10190 CAGGTCTTCTCAGCGCGAGGATCTCAGTCAGCAGTGTCTGTCAGCACCGGCCACCC 10249
 Qy 592 CTCTGTGAACCCAGCAGTCTTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGT 651
 Db 10250 CTCTGCGAGCCAGCTGTCTCTGCTGTGGAGCGGAGTACCTAAAGTCCCTTTGATAAGCT 10309
 Qy 652 CAAGTGCAGAACAGAACCTCCCAAGGAAAGAAATTCGAAGGA---AAATGAATTTAGCTGT 708
 Db 10310 CTGGAGCCACAGAGCCAGCATTTGGAAGAGAGTGGCGAGACCCGAGAGGTTGAGCTGT 10369
 Qy 709 GAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGTGATTTGAGATCCACATGAGAACAC 768
 Db 10370 GATGTGTGGGCAGACATTTCCAGTGTGCTTTGATTTGAGAGCCACATGAAGAAGCAT 10429
 Qy 769 AAGATTTCTTTCACATTACGGGTGAACATGTGCGGAAGAAAGATTCAGGAGCCCTTGGTTT 828
 Db 10430 AAGGACTCTTCACGTATGAGTGTGAGCATGTGCGGGAGGAGATTCAAGGAGCCGTGGTTC 10489
 Qy 829 CTTAAAAATCACATCGGACRCATATGCGAAATCGGGGGCCAGAACGAACTGCAGCAA 888
 Db 10490 CTGAAGAACCATGTCGGACACACATGCGCAAGTCTGSCACAGAGCAGCTTCAGCAA 10549
 Qy 889 GCTTTGGAGATGATCCAGCAACGATCAACGAGGTGCTCCAGGTGCAACGCGCGCAGAGC 948
 Db 10550 GGCATGGAG---AGTCCAGTCACCATCAATGAAGTGTCCAGCCGACGCGCCCTGGGAGC 10606
 Qy 949 ATCTCCTCTCCTTGCAAAATCTGCATGCTTTGTGGCTTCCTATTTCGCAATTAAGAAAGT 1008
 Db 10607 ATCTCACGCGCCTCAAGATCTGCATGCTCTGCGGCTTCCTCTTCCCAATTAAGCAGAGC 10666
 Qy 1009 CTAATTGAGCAGCGCAAGGTGCACACCAAAAGTCTTTGCGTACCAGCAGCGCGCAG 1068
 Db 10667 CTCATTGAGCAGCAGAGGTTCCGCCAAGAACTGTCCCGAGTGCAGCAACGTTGCC 10726
 Qy 1069 ACAGACTCTCCAAAGGAGGAATGCCGCTCTCGAGGGAGGACTTCCCTCAGTGTGTTCAAC 1128
 Db 10727 CTTGATGACCCAGAGAGGAAACCCACGTCCTCCGAGGGAAGAGTTGCTGCAGTTTTTGAAC 10786
 Qy 1129 TTGAGACCAAAATCTCACCTGAAACGGGGAAGAGCCCTGTGCAGATGCTATCCCTCAGCTC 1188
 Db 10787 TTGAGACCCAGATCAACTGTCAGGTAGTACAGTGAAGCCCATGACTGCTATACCTCAGCTT 10846
 Qy 1189 GATCCGTTTCAACACCTTCCAGCTTTGGCAKCTGGCTACCAAGGAAGATGTTGCCATTTTGC 1248
 Db 10847 GACCGTTTCAACACCTACAGGCATGCGCAGTTGGCTACCAAGGAAGAGTGGCGGCTGCC 10906
 Qy 1249 C---AAGAGTGAAGGAATTTGGGCAAGAGGAGCACCAGAACGAGATTCGAGTGTCC 1305
 Db 10907 CAGGAAGAGGTGAAGAGTTCAGGCCAAGAGGAGCAGACAATGACGACTCATGCTCA 10966

Qy 1306 GAGAGGAGCTTGGAG----- 1321
 Db 10967 GAGAAAGAGGAACTAAGGGAATAATGGGTTGGGGTAAGCGGAAGGGTCTGGAAGATCC 11026
 Qy 1322 -AAACAAATAAGAACCAATTGTGTCAGGCTCTCGCAAGAGAAAGAGAGTGAAGCACTCC 1380
 Db 11027 AAACNAGTAAAGCAGTTGTCCAGGCTCTCTCCAGACAAGAGAGCCTAGACATGCT 11086
 Qy 1381 CACGGCAAGCGCCCTCCGTGGAAGCGGATCCCAAGTTTACCAGTGAAGAGAGAGCC 1440
 Db 11087 AATAGTGAAGTGCCTTCTGGGGATAGTACCCCAAGTTGTCCAGTAGCAAGAGAGGCC 11146
 Qy 1441 ACTCAGCTCTCGAGTGGCGCAAGCTTTTCAGAACCTTACCACCACTGCTTGTGCACTCC 1500
 Db 11147 AGCAGCTTCTGAGTGCAGCAAGCCTTCAGGACATACCACAGCTCGTCTGCACTCG 11206
 Qy 1501 AGGGTCC 1507
 Db 11207 AGGGTGC 11213
 RESULT 8
 US-10-087-192-1706
 ; Sequence 1706, Application US/10087192
 ; Publication No. US20020182586A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David W.
 ; APPLICANT: Engelhardt, Eric K.
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
 ; FILE REFERENCE: 529452000122
 ; CURRENT APPLICATION NUMBER: US/10/087,192
 ; CURRENT FILING DATE: 2002-03-01
 ; PRIOR APPLICATION NUMBER: US 09/747,377
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: US 09/798,586
 ; PRIOR FILING DATE: 2001-03-02
 ; NUMBER OF SEQ ID NOS: 2059
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1706
 ; LENGTH: 3016
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 US-10-087-192-1706
 Query Match 35.6%; Score 537; DB 13; Length 3016;
 Best Local Similarity 68.6%; Pred. No. 1.8e-157;
 Matches 834; Conservative 3; Mismatches 324; Indels 55; Gaps 5;
 Qy 344 ATGCCAACTCAATCCCTCTTAATGTACATGGA-TGGCCAAAGAGTGAATGGCAGCTCTCTT 402
 Db 1 ATGCGGACCCAGTCCCTCTCGTGTACATGGAACGCGCGCGGAAGTCTTCAGCAGCTCTCTTA 60
 Qy 403 GCCAGTCCGATGGAGATGGAGATGCTTGTCAATGAAGGGCCNCTGTGTGTCATATTCG 462
 Db 61 GCCTCCAGATGGAGGTGGATGCTGTGCCATAAAGGGCCGCGTGGCAGTCCCCCTC 120
 Qy 463 AGCTACACAAAGAAATAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATTGCAATG 522
 Db 121 CGAGCTGTCTAGGAGAGAGCATGCGCCGTGGCAGAGGGGCCACATGCCCTTGGATTGCAATG 180
 Qy 523 TTCTGCGAGCAGACTTCACACATTCAGAACCTTAATAACATGTCTTAATGCAACAC 582
 Db 181 TTCTGCGAGCCAGGCTTCTCTCAGCGGAGGATCTCAGTCAGCAGCTGCTGCTGCGAGC 240
 Qy 583 CGGCTACCTCTGTGAAACCCAGCAGTTCCTCGGGTTGAAGCAGAGTATCTCAGTCCGCTT 642
 Db 241 CGGCCCACTCTCGAGCCAGCTGTCTGCTGTGGAGCGCGAGTACCTAAGTCCCTT 300
 Qy 643 GATAAAGTCAAGTGCAGAACAGACCTCCCAAGGAAAGAAATTCGAAGA---AAATGA 699
 Db 301 GATAAAGTCTGGAGCCCAAGAGCAGCATTTGGAGAAAGAGTGGCGGAAGACCCCGAGGAG 360

```
QY 700 TTATGCTGTGAGTATGTGGGAGACATTTAGAGTCGCTTTTGTATGTTGAGATCCACATG 759
|||
Db 361 TTGAGCTGTGATGTGTGGGAGACATTTCCAGTGGCTTTTGTATGTTGAGAGCCACATG 420
|||
QY 760 AGAACACACAAAGATCTTTTCACTTACGGGTGTAAATGTCGGGAGAGATTCAGAG 819
|||
Db 421 AAGAAGCATAGGACTCTTCACTATGATGGTGCAGATGTGGGGAGGAGATTCAGAG 480
|||
QY 820 CCTTGTGTTTCTTAAATACATGCGACRCAATATGCAAAATCGGGGCCAGAGCAA 879
|||
Db 481 CGTGTGTTCTGAGNACCATGCGGACACATATGGCACTGTGGCACAGGAGCAAG 540
|||
QY 880 CTGACAGAGGCTGTGAGATGTAGTCAGCAACGATCAACAGAGTGTGTGAGTGCACGCG 939
|||
Db 541 CTTTCAAGAGGAGTGGAG---AGTCAGTCAACATCAATGAAGTGTGTGAGCGGACGCG 597
|||
QY 940 GCGGAGAGCATCTCTCTCTTGCATAATCTGCATGTTTGTGGTCTCTATTTCCAAAT 999
|||
Db 598 CTGGGAGCATCTCCACGCCCTTACAGATCTGCATGTTGTGGGCTTCTCTTCCCAAT 657
|||
QY 1000 AAAGAAGTCTAATTCAGCACCGCAGAGTGCACCAAAAATCTGCTTTCGGTACCAGC 1059
|||
Db 658 AAGCAGAGCTCATTTGAGCAGAGAGGTTTACGCCAAGAAACTGTCCCAAGTGCAGC 717
|||
QY 1060 AGCGGCGAGACAGACTCTCCACAAAGAGGAATGCCGTCTTCAGGGAGGACTTCTGTGCAG 1119
|||
Db 718 AACGTTGCCCTGTATGATCCACCGAGAGGAACCCAGTCCCGGAGGAGTGTGTGCAG 777
|||
QY 1120 TTGTTCAACTTCAGACCAAAATCTACCTGTAACCGGGAGAGCTGTTCAGATGCATC 1179
|||
Db 778 TTTTGAATTCAGACCCCAAGATCAACTGCAGGTAGTACAGTGAAGCCCATGACCTGCATA 837
|||
QY 1180 CCTCAGCTCGATCCGTTTCAACACTTCCAGGCTTGCACACTGCTACCAAGGAAGAGTT 1239
|||
Db 838 CTTAGCTTTGACCCGTTTACCACTTACAGGCTGTCAGTGGTGTGCTACCAAGGAAGAGT 897
|||
QY 1240 GCCATTTGCC---AAGAAGTGAAGAAATTTGGGGCAAGAGGAGGAGCACCAACGACGAT 1296
|||
Db 898 GCCGTTGCCAGGAGAGAGTGAAGAGTTCAGGCCAAGGAAGGACACACAAATGACGAC 957
|||
QY 1297 TCGAGTTCGAGAGAGGCTTGGAG----- 1321
|||
Db 958 TCATGCTCAGAGAGAGAGAACTATGGGGAAATATGGGTTGGGGTAAAGCGGAGGCT 1017
|||
QY 1322 -----AAACAAATGAAGAACCATTTGTGAGGCTCTCCAGAGAGAGAGAGTGC 1371
|||
Db 1018 GGAAGTCCAAACAAAGTAAAGCAGTGTGTCCAGGTCTCTCCCAAGACAAAGAGAGCCT 1077
|||
QY 1372 AAACACTCCAGCGGAGAGCGCCCTCCGTGAGAGCGGATCCCAAGTTACCCAGTAGCAAG 1431
|||
Db 1078 AGACATGCTAATAGTGAAGTGCCTTCTGGGATAGTGACCCCAAGTTGTCCAGTAGCAAG 1137
|||
QY 1432 GAGAAGCCCACTCATCTGTCCAGTGCAGGCAAGCTTTAGAACCTTACACAGCTGTGTC 1491
|||
Db 1138 GAGAAGCCCACTCATCTGTGAGTGCAGAAAGCTTTCAGGACATACCAAGCTGTGTC 1197
|||
QY 1492 TTGCACTCCAGGCTCC 1507
|||
Db 1198 CTGCACTCGAGGCTC 1213
|||

RESULT 9
US-10-029-386-9711/c
; Sequence 9711, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
```

```
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 9711
; LENGTH: 530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR20.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.94
; OTHER INFORMATION: SWISSPROT HIT: O75362, EVALUE 2.00e-67
; OTHER INFORMATION: NT HIT: g11421959, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BF237807.1, EVALUE 0.00e+00
US-10-029-386-9711

Query Match 14.3%; Score 215.6; DB 16; Length 530;
Best Local Similarity 98.2%; Pred No. 1.2e-56;
Matches 218; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1286 ACAACGACGATTCGAGTTCGAGAGAGGAGCTTTGGAGAAACAAATAAGAACCACTTGTGCAG 1345
|||
Db 530 ACAACGACGATTCGAGTTCGAGAGAGGAGCTTTGGAGAAACAAATAAGGCGAGTGTGTGCAG 471
|||
QY 1346 GCCTCTCGCAAGAGAAAGAGAGTGCACACACTCCACGGCGAAGCGCCCTCCGTGGAGC 1405
|||
Db 470 GCCTCTCGCAAGAGAAAGAGAGTGCACACACTCCACGGCGAAGCGCCCTCCGTGGAGC 411
|||
QY 1406 CGGATCCCAAGTTACCCAGTAGCAAGGAGAGCCCACTCCTCCGAGTCCGAGTCGCGCAAG 1465
|||
Db 410 CGGATCCCAAGTTACCCAGTAGCAAGGAGAGCCCACTCCTCCGAGTCCGAGTCGCGCAAG 351
|||
QY 1466 CTTTCAGAACCTTACCACGAGCTGTGCTTTCGACTCCAGGGTCC 1507
|||
Db 350 CTTTCAGAACCTTACCACGAGCTGTGCTTTCGACTCCAGGGTCC 309
|||

RESULT 10
US-10-040-739-520
; Sequence 520, Application US/10040739
; Publication No. US20020173635A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; McCoy, John
; LaVallie, Edward
; Racie, Lisa
; Merberg, David
; Treacy, Maurice
; Spaulding, Vikki
; TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
; NUMBER OF SEQUENCES: 1519
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/040,739
; FILING DATE: 07-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
```

```

; APPLICATION NUMBER: 09/036,520
; FILING DATE: 03-JUN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 520:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 469 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 520:
US-10-040-739-520

Query Match          10.4%; Score 156.4; DB 13; Length 469;
Best Local Similarity 99.4%; Pred. No. 5.2e-38;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1350 CTCGCAAGAGAAAGAGAGTGAACACTCCACACGGCGAAGCGCCCTCGTGGAGCGCGA 1409
DB 19 CTAGCAAGAGAAAGAGAGTGAACACTCCACACGGCGAAGCGCCCTCGTGGAGCGCGA 78
QY 1410 TCCCAAGTTACCCAGTAGCAGGAGAGCGCACTCACTGCTCCGAGTGGCGCAAGCTTT 1469
DB 79 TCCCAAGTTACCCAGTAGCAGGAGAGCGCACTCACTGCTCCGAGTGGCGCAAGCTTT 138
QY 1470 CAGAACCTTACCACCACTGCTGTTGCTCACTTCCAGGGTCC 1507
DB 139 CAGAACCTTACCACCACTGCTGTTGCTCACTTCCAGGGTCC 176

RESULT 11
US-10-172-118-1511
; Sequence 1511, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1511
; LENGTH: 6033
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM 014643
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1511

Query Match          6.8%; Score 103.2; DB 17; Length 6033;
Best Local Similarity 52.7%; Pred. No. 1.3e-20;
Matches 291; Conservative 3; Mismatches 246; Indels 12; Gaps 3;

QY 697 GAATTACTGTGAGGTATGTGGGAGACATTTAGATCGCTTTTGATGTTGAGATCCAC 756
DB 1057 GAGTTCCTCGGAGGTGTGTGGCCAGGCTTTCAGCCAGACCTGTGTTCTGAAGCGGCAC 1116
QY 757 ATGAGAACACAAAGATTTCTTCTACTTACGGGTGTAAACATGTGCGGAAGAAGATTCAAG 816

; APPLICATION NUMBER: 09/036,520
; FILING DATE: 03-JUN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 520:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 469 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 520:
US-10-040-739-520

Query Match          10.4%; Score 156.4; DB 13; Length 469;
Best Local Similarity 99.4%; Pred. No. 5.2e-38;
Matches 157; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1350 CTCGCAAGAGAAAGAGAGTGAACACTCCACACGGCGAAGCGCCCTCGTGGAGCGCGA 1409
DB 19 CTAGCAAGAGAAAGAGAGTGAACACTCCACACGGCGAAGCGCCCTCGTGGAGCGCGA 78
QY 1410 TCCCAAGTTACCCAGTAGCAGGAGAGCGCACTCACTGCTCCGAGTGGCGCAAGCTTT 1469
DB 79 TCCCAAGTTACCCAGTAGCAGGAGAGCGCACTCACTGCTCCGAGTGGCGCAAGCTTT 138
QY 1470 CAGAACCTTACCACCACTGCTGTTGCTCACTTCCAGGGTCC 1507
DB 139 CAGAACCTTACCACCACTGCTGTTGCTCACTTCCAGGGTCC 176

RESULT 12
US-10-342-887-1511
; Sequence 1511, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1511
; LENGTH: 6033
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1511

Query Match          6.8%; Score 103.2; DB 18; Length 6033;
Best Local Similarity 52.7%; Pred. No. 1.3e-20;
Matches 291; Conservative 3; Mismatches 246; Indels 12; Gaps 3;

QY 697 GAATTACTGTGAGGTATGTGGGAGACATTTAGATCGCTTTTGATGTTGAGATCCAC 756
DB 1057 GAGTTCCTCGGAGGTGTGTGGCCAGGCTTTCAGCCAGACCTGTGTTCTGAAGCGGCAC 1116
```

Qy	757	ATGAGAACACACAAGATCTCTTTTCACTTACGGGTGTAACTATGTCGGGAAGAAGATTCAAG	816
Db	1117	ATGAAGAAGACACCGGGCTCTTCGACCA CGGCTGCCACATCTCGCGCGTAGGTTCAAG	1176
Qy	817	GAGCCTTGGTTTCTTAAAAATCATATCGGACNCRATATGCGCAATTCGGGGCCAGAAGC	876
Db	1177	GAGCCCTGGTTCTCTAAGAACCACATGAAGGCGCACGGCCCCCAAGACGGGCAGCAAGAAC	1236
Qy	877	AAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCTGTCAGGTGCAC	936
Db	1237	AGGCCCAAGAGTAGCTGGACCCCA---TCGCCACCATCAACACGTGTCTCAGAGAGAG	1293
Qy	937	CGCGCCGAGAGCATCTCCTCTCTTGC AAAATCTGSCATGGTTTGTGGCTTCCTATTTTCCA	996
Db	1294	GTGATCGTTCGCGGCTCTGAGCCTCTACGAGTCTTCGCCAAGTCGCGGAACCTGTTTACA	1353
Qy	997	AATTAAGAAGTCTAATTGAGACCGCAAGGTGCACACCAAAAAAATGTCTTCGGTAGCC	1056
Db	1354	AACCTGCAGAGCTTGAACGCGCCACAATGCCATCCACCGCAGATCGAGGC---CAGCGCG	1410
Qy	1057	AGCAGCGGCAGACAGACTCTCCACAGGAGGAATGCGGTCTTCGAGGGAGGACTTCTCTG	1116
Db	1411	ACGCGCGCCCGCGCGAGAGGGGGCGGCGCCCTCGGACACCAAGCAGTTCCTTTCTC	1470
Qy	1117	CAGTTGTTTCAACTTGAGACCAAAATCTCACCCCTCAAAACG-----GGGAAGAAGCCGTGTC	1170
Db	1471	CAGTGCCTGAACCTGAGGCGGTTCGGCGCGCGGCGACTCTGTGCCCTTGCAACGAGGCCGGA	1530
Qy	1171	AGATGCATCCCTCAGCTCGATCCGTTTCAACCACTTCCAGGGTTGGCAKCTGGCTACCAAA	1230
Db	1531	CGGCGGGTGGCTGAGCTGGACCCGGTCAACAGCTTACCAGGCTTCGACCTGGCAGCTGGCCACGCGG	1590
Qy	1231	GGAAWAGTTGCC	1242
Db	1591	GGTAAGGTGGCC	1602

```

RESULT 13
US-10-074-475-93
; Sequence 93, Application US/10074475
; Publication No. US20030092898A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Hu, Ping
; APPLICANT: Recipon, Herve
; APPLICANT: Karra, Kalpana
; APPLICANT: Cafferkey, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
; TITLE OF INVENTION: Genes and Proteins
; FILE REFERENCE: DEX-0313
; CURRENT APPLICATION NUMBER: US/10/074,475
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: 60/268,292
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 295
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 93
; LENGTH: 8156
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-074-475-93

```

Qy	757	ATGAGAACACACAAAGATCTCTTTCACTTACGGGTGTAAACATGTGCCGAAGAAGATTCAAG	816
Db	1117	ATGAAGAAGCACCGGGCTCTTTTGGACACCGGTGGCCACATCTGCGGCCGTAGTTTCAAG	1176
Qy	817	GAGCCTTGGTTTCTTAAATATCATCTCGGACACATAATGGCAATTCGGGGGCCAAGCC	876
Db	1177	GAGCCCTGGTTCTTCAAGAACCATGAAGGGCGCACGGCCCCCAAGACCGGCACGCAAGAAC	1236
Qy	877	AAATCTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCTGTCACAGGTGTCAC	936
Db	1237	AGGCCCAAGAGTGAGCTGGAACCCCA---TCGCCACCATCAACACGTGTCTCAGAGGAG	1293
Qy	937	CGCGCCGAGAGCATCTCTCTCTTGCAAAATCTGCATGGTTTGTGGCTTCTCTATTTCGA	996
Db	1294	GTGATCGTCCCGGCTGAGCCTCTACGAGGTCTGCGCCAAGTTCGGGAACCTGTTTACA	1353
Qy	997	AATAAAGAAAGTCTAATTGAGCAACCGCAAGGTGCACACCAAAAAAATCTGTTTCGGTACC	1056
Db	1354	AACCTGCACAGCTTGAACGCCCAATGCCATCCAACGACAGAGTCGAGGC---CAGCCGC	1410
Qy	1057	AGCAGCGCGACAGACAGACTCTCCACAAGGAGGAATGCGCTCTCGAGGGGAGGACTTCCTG	1116
Db	1411	ACGCGCGCCCGCCGAGGAGGGGGCGGAGGGGCCCTCGGACACCAAGCAGTCTCTTCTC	1470
Qy	1117	CAGTTGTTTCAACTTGAGACCCAAATCTCACCTCTGAACG-----GGGAAGAAGACCTGTC	1170
Db	1471	CAGTGCTTGAACCTGAGCGCGTTCGGGGCGCGGCGACTCTGTGCCCTGGCACGCAAGGCGGA	1530
Qy	1171	AGATGCATCCCTCAGCTCGATCCGTTTCAACACCTTCCAGGCTTGGCAKCTGGCTACCAAA	1230
Db	1531	CGCGGGTGGCTGAGCTGGACCGCGGTCAACAGACTACGAGGCTGGCAGCTGGCCACGCGG	1590
Qy	1231	GGAAWAGTTGCC	1242
Db	1591	GGTAAGTGGCC	1602

```

RESULT 14
US-10-029-386-23411/c
; Sequence 23411, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USED
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 23411
; LENGTH: 251
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR20.3
; OTHER INFORMATION: EXPRESSED IN LONG, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.94
; OTHER INFORMATION: EST_HUMAN HIT: W05407.1, EVALUATE 1.00e-117
; OTHER INFORMATION: NT HIT: G111421959, EVALUATE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: O75362, EVALUATE 1.00e-38
US-10-029-386-23411

```

```
Query Match      6.4%; Score 97; DB 16; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.9e-19;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	697	GAATTTAGCTGCGAGGTATGTGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCAC	756
Db	1057	GAGTTCCTCCGTCGAGGTGTGTGCGCAGCGCTTCAGCCAGACCTCGTGTCTCGAAGGCGCAC	1116

QY 1411 CCCAGTTACCCAGTAGCAAGAGAGCCCACTCACTGCTCCGAGTGGCGCAAGCTTTTC 1470
 |||||
 Db 251 CCCAAGTTACCCAGTAGCAAGAGAGCCCACTCACTGCTCCGAGTGGCGCAAGCTTTTC 192
 |||||

QY 1471 AGAACCTACCAACAGCTGGTCTTTCACCTCCAGGGTCC 1507
 |||||
 Db 191 AGAACCTACCAACAGCTGGTCTTTCACCTCCAGGGTCC 155
 |||||

RESULT 15

US-10-758-307-110
 ; Sequence 110, Application US/10758307
 ; Publication No. US20040209290A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GENOMIC HEALTH, INC.
 ; APPLICANT: RUSH UNIVERSITY MEDICAL CENTER
 ; APPLICANT: Cobleigh, Melody
 ; APPLICANT: Shak, Steven
 ; APPLICANT: Baker, Joffre
 ; APPLICANT: Cronin, Maureen
 ; TITLE OF INVENTION: GENE EXPRESSION MARKERS FOR BREAST
 ; FILE REFERENCE: 39740/0008 US
 ; CURRENT APPLICATION NUMBER: US/10/758,307
 ; CURRENT FILING DATE: 2004-01-14
 ; PRIOR APPLICATION NUMBER: US 60/440,861
 ; PRIOR FILING DATE: 2003-01-15
 ; NUMBER OF SEQ ID NOS: 440
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 110
 ; LENGTH: 70
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Amplicon
 US-10-758-307-110

Query Match 4.6%; Score 70; DB 20; Length 70;
 Best Local Similarity 100.0%; Pred.No. 2.9e-11;
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1419 ACCCAGTAGCAAGAGAGCCCACTCACTGCTCCGAGTGGCGCAAGCTTTTCAGAACCTA 1478
 |||||
 Db 1 ACCCAGTAGCAAGAGAGCCCACTCACTGCTCCGAGTGGCGCAAGCTTTTCAGAACCTA 60
 |||||

QY 1479 CCACCAGCTG 1488
 |||||
 Db 61 CCACCAGCTG 70
 |||||

Search completed: June 29, 2005, 12:01:53
 Job time : 593.497 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 15:06:07 ; Search time 3243.1 Seconds
(without alignments)
17687.639 Million cell updates/sec

Title: US-08-731-499-3

Perfect score: 1507

Sequence: 1 GCAGTGTCTGGATTGACT.....GGCTTGCATCCAGGGTCC 1507

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	590.4	39.2	595	4	BF944766	BF944766 PM3-NN117
2	557.4	37.0	618	5	EX955300	EX955300 DKF2781N
3	529.6	35.1	640	4	BM834501	BM834501 K-EST0109
4	436.6	29.0	571	7	CR753532	CR753532 DKF2p469L
5	434.4	28.8	895	9	CR094054	CR094054 Forward s
6	434.2	28.8	569	2	AW753609	AW753609 RCL1-CT026
7	401	26.6	585	5	BP366185	BP366185 BP366185
8	394.4	26.2	598	5	BP270262	BP270262 BP270262
9	389	25.8	583	5	BP229047	BP229047 BP229047
10	376.2	25.0	636	6	CB457791	CB457791 715904 MA
11	359.2	23.8	594	7	CV409050	CV409050 RCL1-HT026
12	343.8	22.8	662	6	CB460210	CB460210 720064 MA
13	336	22.3	458	5	BP201868	BP201868 BP201868
14	302	20.0	352	4	BF963063	BF963063 PM3-NN117
15	289.6	19.2	768	2	BP237807	BP237807 601842326
16	285.6	19.0	412	1	AA069107	AA069107 zml0b10.r
17	262.8	17.4	313	2	BF803316	BF803316 RCL1-C1011
18	243.4	16.2	307	8	AZ537566	AZ537566 AST-2P596
19	226	15.0	226	1	AA451924	AA451924 aa85e10.s
20	224	14.9	304	2	BF803309	BF803309 RCL1-C1011
21	221.8	14.7	553	1	AJ743841	AJ743841 AJ743841
22	215	14.3	590	6	CD689109	CD689109 EST5631 h
23	201.6	13.4	567	1	A1552996	A1552996 ve41c05.y
24	201.4	13.4	885	7	CR432863	CR432863 CR432863

ALIGNMENTS

RESULT 1
BF944766 595 bp mRNA linear EST 22-JAN-2001
LOCUS PM3-NN1174-181000-001-c08 NN1174 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF944766
ACCESSION BF944766.1 GI:12362041
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 595)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,P.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?c1=PM3&t2=PM3-NN1174-181000-001-c08&t3=2000-10-18&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 585.
Location/Qualifiers
1..595
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="NN1174"
/note="Organ: nervous normal; Vector: puc18; Site 1: Sma1; Site 2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application

25 178.6 11.9 884 6 CA789554
26 160.2 10.6 477 6 CB548508
27 159 10.6 915 6 CA981247
28 140 9.3 647 1 AL595898
29 138.4 9.2 818 9 CR127300
30 137.4 9.1 376 1 AA646529
31 132.8 8.8 735 5 BP675362
32 130.8 8.7 505 1 A1467614
33 130.2 8.6 603 4 BJ031636
34 122.4 8.1 966 5 BU912448
35 118.8 7.9 692 2 BB617224
36 117.2 7.8 2331 3 AK030249
37 116.2 7.7 672 5 BX854767
38 114.8 7.6 342 5 BY206008
39 114.8 7.6 361 5 BY207082
40 114.8 7.6 377 5 BY200588
41 111 7.4 338 4 BG083784
42 108.2 7.2 133 2 BE706424
43 104.4 6.9 793 6 CA752347
44 104 6.9 805 6 CB522302
45 96.2 6.4 295 1 AJ743842

CA789554 AGENCOURT
CB548508 AMGNNUC:M
CA981247 AGENCOURT
AL595898 AL595898
CR127300 Forward s
AA646529 ve41c05.y
BP675362 BP675362
A1467614 ve41c05.x
BJ031636 BJ031636
BU912448 AGENCOURT
BB617224 BB617224
AK030249 MUR MURCU
BX854767 BX854767
BY206008 BY206008
BY207082 BY207082
BY200588 BY200588
BG083784 H3092A10-
BE706424 RCL1-HT026
CA752347 UI-M-FOO-
CB522302 UI-M-GHO-
AJ743842 AJ743842

No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```
ORIGIN
Query Match      39.2%; Score 590.4; DB 4; Length 595;
Best Local Similarity 99.5%; Pred. No. 3.3e-164;
Matches 591; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 615 GGTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAGACAGAACCTCCCAA 674
Db 1 GGTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAGACAGAACCTCCCAA 60

QY 675 GGAAGAATTCGAGGAAATGAATTTAGCTGCTGAGGTATGTGGCGACACATTTAGAGT 734
Db 61 GGAAGAATTCGAGGAAATGAATTTAGCTGCTGAGGTATGTGGCGACACATTTAGAGT 120

QY 735 CGCTTTTGATGTGAGATCCACATGAGAACACACAAAAGATCTTTTCACTTACGGGTGAA 794
Db 121 CGCTTTTGATGTGAGATCCACATGAGAACACACAAAAGATCTTTTCACTTACGGGTGAA 180

QY 795 CATGCGGGAAGAAGATCAAGAGCCTTGGTTCTTAAATAATCATGCGGACACATAA 854
Db 181 CATGCGGGAAGAAGATCAAGAGCCTTGGTTCTTAAATAATCATGCGGACACATAA 240

QY 855 TGGCAATCGGGGGCCAGNAGAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACCAT 914
Db 241 TGGCAATCGGGGGCCAGNAGAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACCAT 300

QY 915 CAACGAGGTGCTCAGGTGACGCGCGGAGAGCATCTCTCTCTGTCMAAATCTGCAT 974
Db 301 CAACGAGGTGCTCAGGTGACGCGCGGAGAGCATCTCTCTCTGTCMAAATCTGCAT 360

QY 975 GGTGTGCGTCTCTATTTCCTTCAAAATAAAGAAAGTCTAATTTAGAGCCGCAAGGTGCAC 1034
Db 361 GGTGTGCGTCTCTATTTCCTTCAAAATAAAGAAAGTCTAATTTAGAGCCGCAAGGTGCAC 420

QY 1035 CAAAAAACTGTTTCGGTACACGAGCGCGGAGAGCATCTTCCACAGGAGGATGCC 1094
Db 421 CAAAAAACTGTTTCGGTACACGAGCGCGGAGAGCATCTTCCACAGGAGGATGCC 480

QY 1095 GTCCCTCGAGGAGGACTTCTCAGTGTGTTCAACTTGAGACCAAAATCTCACCCCTGAAAC 1154
Db 481 GTCCCTCGAGGAGGACTTCTCAGTGTGTTCAACTTGAGACCAAAATCTCACCCCTGAAAC 540

QY 1155 GGGGAAGAAGCTGTGAGATGATCCCTCAGTCTGATCGTTCAGTCCACCTTCCA 1208
Db 541 GGGGAAGAAGCTGTGAGATGATCCCTCAGTCTGATCGTTCAGTCCACCTTCCA 594
```

```
RESULT 2
BX955300
LOCUS
DEFINITION
  DKFZp781N0270 r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
ACCESSION
  BX955300
VERSION
  BX955300.1 GI:43437635
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 618)
  Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
  Fobo,G., Han,M. and Wiemann,S.
  EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., et al.)
  Unpublished (2003)
  Contact: MIPS
  MIPS
```

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No sl sequence available.
This clone (DKFZp781N0270) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

```
FEATURES
Location/Qualifiers
  1..618
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="DKFZp781N0270"
    /dev_stage="adult"
    /lab_host="DH10B"
    /clone_lib="781 (synonym: hlcc4)"
    /note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
    cDNA-collection"
```

ORIGIN

```
Query Match      37.0%; Score 557.4; DB 5; Length 618;
Best Local Similarity 95.8%; Pred. No. 2.3e-154;
Matches 592; Conservative 1; Mismatches 23; Indels 2; Gaps 2;

QY 377 GGCCAAAGAGTGAATGGCAGCTCTCTTGCAGTCCGATGGAGATGGA-GATGCTTTGTCAA 435
Db 2 GGCCAAAGAGTGAATGGCAGCTCTCTTGCAGTCCGATGGAGATGGAAGATGCTTTGTCAA 61

QY 436 TGAAGAGGCGCCNCTGTCTCAATTCGAGCTACACAAAGAAAAAATGTCAATCCGAATC 495
Db 62 TGAAGAGGCGCCGCTGTGTTCATTCGAGCTACACAAAGAAAAAATGTG-ATCCAAATC 120

QY 496 GAGGGGAATATGCCCTTGGATTCATGTTCTGCAGCAGACCTTCACATTCAGAAC 555
Db 121 GAGGGGAATATGCCCTTGGATTCATGTTCTGCAGCAGACCTTCACATTCAGAAC 180

QY 556 CTTAATAAACATGTCTTAATGCAACACCGGCTACCCCTCTGTGAACAGCAGATCTTCGG 615
Db 191 CTTAATAAACATGTCTTAATGCAACACCGGCTACCCCTCTGTGAACAGCAGATCTTCGG 240

QY 616 GTTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAGAACCACTCCCAAG 675
Db 241 GTTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAGAACCACTCCCAAG 300

QY 676 GAAAGATTCAGGAAATGAATTTAGCTGTGAGTATGTGGCAGACATTTAGAGTC 735
Db 301 GAAAGATTCAGGAAATGAATTTAGCTGTGAGTATGTGGCAGACATTTAGAGTC 360

QY 736 GCTTTTGATGTGAGATCCACATGAGAACACACAAAGATCTTTTCACTTACGGGTGAAC 795
Db 361 GCTTTTGATGTGAGATCCACATGAGAACACACAAAGATCTTTTCACTTACGGGTGAAC 420

QY 796 ATGTGCGGAAGAAGATTCAGAGCCTTGTTCTTAAAAATTCATGCGGACACATAT 855
Db 421 ATGTGCGGAAGAAGATTCAGAGCCTTGTTCTTAAAAATTCATGCGGACACATAT 480

QY 856 GGCAATCGGGGGCCAGNAGAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATC 915
Db 481 GGCAATCGGGGGCCAGNAGAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATC 540

QY 916 AACGAGGTGCTCCAGGTGCGCGGCCGAGAGCATCTCTCTCTCTTAAAAATTCGATG 975
Db 541 AACGAGGTGCTCCAGGTGCGCGGCCGAGAGCATCTCTCTCTCTTAAAAATTCGATG 600

QY 976 GTTTGTGGCTTCCATTT 993
Db 601 GTTTGTGGCTTCCATTT 618
```

RESULT 3

BM834501

LOCUS

DEFINITION

BM834501 640 bp mRNA linear EST 06-MAR-2002
K-EST0109502 S11SNU1 Homo sapiens cDNA clone S11SNU1-63-D04 5',

```

mRNA sequence.
BM834501.1 GI:19190910
EST.
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 640)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 63 row: D column: 04
High quality sequence stop: 640.
Location/Qualifiers
1..640
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S115NUI-63-D04"
/sex="M"
/tissue_type="Stomach"
/cell_type="Lymphoblast-like"
/lab_host="SNU-1"
/clone_lib="S115NUI"
/note="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI;
Site 2: XhoI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including SfiI
site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized with Superscript II using SfiI
oligo-dT primer. After first strand synthesis, RNA was
degraded by NaOH treatment and cDNA was amplified by PCR
reaction. The PCR products were digested with SfiI and
cloned into DraIII- digested pME18S-FL3 vector. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
'full-length enriched cDNA library.'"

FEATURES
source
1..571
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZp469L046"
/tissue_type="kidney"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="469"
/note="Vector: pSport1_Sfi; Site 1: SfiI; Site 2: SfiI"

ORIGIN
Query Match 35.1%; Score 529.6; DB 4; Length 640;
Best Local Similarity 95.2%; Pred. No. 4.6e-146;
Matches 610; Conservative 0; Mismatches 25; Indels 6; Gaps 6;

QY 67 TCGTTTGTGCTCATTTTGAATTCGAGGCAATTTGGGAACCTGTATGCCCTTTGTTGTG 126
DB 1 TCGTTTGTGCTCATTTTGAATTCGAGGCAATTTGGGAACCTGTATGCCCTTTGTTGTG 60

QY 127 GAAAGAACCGAGTGACACCATCACTGAGCTTCTTAAAGTTTGAAGTTAGAGACTAT 186
DB 61 GAAAGAACCGAGTGACACCATCACTGAGCTTCTTAAAGTTTGAAGTTAGAGACTAT 120

QY 187 ACACATTTCTTTTGAACCTTTTATAATAATTTGCTCTGG-TTTTGGACCCGAGCACTG 245
DB 121 ACACATTTCTTTTGAACCTTTTATAATAATTTGCTCTGGTTTGGACCCGAGCACTG 180

QY 246 TAGA-GGGTGAGTGACAGGTCTTAC-AGTGGCTCTTAATCCAACCTCAGAAAATTGCCCAAC 303
DB 181 TAGAGGGGTGAGTGACAGGTCTTACAGTGGCTTATTCCAACTCAGAAATTGCCCAAC 240

QY 304 GGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTT 363

```

```

Db 241 GGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTT 300
QY 364 AATGTACATGGAT-GGCCAAAGAGTATTGGGAGCTCTCTTGGCAGTCCGATCGAGATGGA 422
Db 301 AATGTACATGGATGGCCAAAGAGTATTGGGAGCTCTCTTGGCAGTCCGATCGAGATGGA 360
QY 423 -GATGCCCTTGTCAATGAAGAGGCCCNCTGTGTGTTCAATTCCGAGCTACACAAAGAAAAA 481
Db 361 GGATGCCCTTGTCAATGAAGAGGCCCGCTGTGTGTTCCATTCCGAGCTACACAAAGAAAAA 420
QY 482 TGTCAATCGATCGAGGGGATATGCCCTTGGATTGATGTTCTGCGAGCCAGACCTTCA 541
Db 421 TGTCT-ATCCAAATCGAGGGGTATATGCCCTTGGATTGATGTTCTGCGAGCCAGACCTTCA 479
QY 542 CACATTGAGAGACCTTAATAAACAATGTCTTAATGCAACACCGGCTACCCCTCTGTGAAC 601
Db 480 CACATTGAGAGACCTTAATAAACAATGTCTTAATGCAACACCGGCTACCCCTCTGTGAAC 539
QY 602 CAGCAGTTCTTCGGGTTCAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTCCGAA 661
Db 540 CAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTCCGAA 599
QY 662 CAGACCTTCCCAAGCAAGAAAGATTCAGAGGAAATGAATTT 702
Db 600 CAGACCTTCCCAAGCAAGAAAGATTCAGAGGAAATGAATTT 640

RESULT 4
CR753532 571 bp mRNA linear EST 01-SEP-2004
LOCUS DKFZp469L046_r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone
DEFINITION DKFZp469L046 5', mRNA sequence.
ACCESSION CR753532
VERSION CR753532.1 GI:51845942
KEYWORDS EST.
SOURCE Pongo pygmaeus (orangutan)
ORGANISM Pongo pygmaeus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
1 (bases 1 to 571)
Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
Fobo,G., Han,M. and Wiemann,S.
Pongo pygmaeus mRNA (Wambutt,R., Heubner,D., Mewes,H.W., et al.)
Unpublished (2004)
Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert. Clone from S. Wiemann,
Molecular Genome Analysis, German Cancer Research Center (DKFZ);
Email s.wiemann@dkfz-heidelberg.de; sequenced by Agowa
(Berlin/Germany) within the cDNA sequencing consortium of the
German Genome Project. This clone (DKFZp469L046) is available at
the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in
Berlin, Germany. Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469L046
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.
Location/Qualifiers
1..571
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZp469L046"
/tissue_type="kidney"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="469"
/note="Vector: pSport1_Sfi; Site 1: SfiI; Site 2: SfiI"

ORIGIN
Query Match 29.0%; Score 436.6; DB 7; Length 571;
Best Local Similarity 92.0%; Pred. No. 2.5e-118;

```

Matches 526; Conservative 0; Mismatches 40; Indels 6; Gaps 6;

QY 56 AAGACCACTAATGCTTTGTGCTGATTATATTTGAATCGAGCATTGGAAACCCCTGTATG 115
 Db 1 AAGAGAACTAATGCTTTGTGCTGATTATATTTGAATCAAGGCATTTGGAGCCCTGTATG 60

QY 116 CTTCTTTGTGGAAGAACCCAGTGACACCATCACCTGAGCTTCTTAAAGATTTCGAAGAAGT 175
 Db 61 CCGTGTGTTGGAAGAACCCAGTGACACCATCACCTGAGCTTCTTAAAGATTTCGAAGAAGT 120

QY 176 TAGAGGACTATACACTTTCTTTTGAACCTTTTATATAAATATTTGCTCTGCTTTTGGAA- 234
 Db 121 TAGAGGACTATACGCTTTCTTTTGAACCTTTTATATAAATATTTGCTCTGCTTTTGGAA 180

QY 235 CCCAGGACTGTTAGA- GGGTGAGTGACAGGCTTTAC- AGTGGCCCTTAATCCMACTCCAGA 292
 Db 181 CCCAGGCTGTTAGAGGGGTGAGTGACCAAGTCTTACAAGTGCCCTTATTTCCAACCTCCAGA 240

QY 293 AATTGCCCAACGGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCAACT 352
 Db 241 AATTCCCCAACGGAACCTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCAACT 300

QY 353 CAATCCCTCTTAATGTACATGAT- GGCCAAAGAGTGATGGCAGCTCTCTTGGCCAGTCCG 411
 Db 301 CAGTCCCTCTTAATGTACATGATGGCCAGAGTGATGGCAGCTCTCTTGGCAGTCA 360

QY 412 ATGGAGATGGA- GATGCCCTTGTCAATGAAGGGCCCTGTTGTCATTCGAGGTACAC 470
 Db 361 ATGGAGATGGAGGAGCCCATGTCAATGAAGGGAGCCGCTGTTGTTCCATTCGAGCTACG 420

QY 471 AAAGAAAAAATGTCAATCCGAATCGAGGGGAATATGCCCTTGGATGTCATGTTCTGCGAG 530
 Db 421 CAAGAAAAAATGTCT- ATCCAAATCGAGGGGTATATGCCCTTGGATGTCATGTTCTGCGAG 479

QY 531 CCAGACCTTTCACATTCAGAGACCTTAATAAATCATGTCTTAATGCAACACCGCCCTAC 590
 Db 480 CCAGACCTTTCACATTCAGAGACCTTAATAAATCATGTCTTAATGCAACACCGCCCTAC 539

QY 591 CCTGTGTAACAGCAGTCTTCGGGTTGAAG 622
 Db 540 CCTGTGTAACAGCAGTCTTCGGGTTGAAG 571

RESULT 5
 CR094054
 LOCUS Forward strand read from insert in 5'HPRT insertion targeting and chromosome engineering clone MHPN19116, genomic survey sequence.
 ACCESSION CR094054
 VERSION CR094054.1 GI:49827883
 KEYWORDS GSS; genome survey sequence; MICER.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 895)
 AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.
 TITLE Direct Submission
 JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. http://www.sanger.ac.uk/MICER
 FEATURES Location/Qualifiers
 source 1..895
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="MHPN19116"
 /clone_lib="MHPN"

ORIGIN

Query Match 28.8%; Score 434.4; DB 9; Length 895;
 Best Local Similarity 69.0%; Pred. No. 1.3e-117;

Matches 620; Conservative 2; Mismatches 270; Indels 6; Gaps 2;

QY 332 GTGACAGGAAACATGCGCAACTCAATCCCTCTTAATGTACATGGATGGCCAGAGTGATTG 391
 Db 1 GTGGTTGGCAGCATGCCGACCCAGTCCCTCTCGTGTCTATGACGCGCGGAAGTCTCTCA 60

QY 392 GCAGCTCTCTTGGCCAGTCCGATGAGATGGAGATGCCCTTGTCAATGAAGGCCCNCTGT 451
 Db 61 GCAGCTCTCTAGGTCCCGATGAGATGGAGTGTGATGTCTGTGCCCATAAAGGCCCGGTGG 120

QY 452 TGTCAATTCGAGCTTACACAAAGAAAAAATGTCAATCCGAATCGAGGGGAATATGCCCT 511
 Db 121 CAGTCCCTCTTCGAGCTGCTCAGGAGAGAGCATGGCGTGGCAGAGGCCACATGCCCC 180

QY 512 TGGATTGATGTTCTGAGCCAGACCTTTCACATTCAGAAGACCTTAATAACATGTCT 571
 Db 181 TGGATTGATGTTCTGAGCCAGGCTCTCTCTCAGGGCGAGGATCTCTAGTCAGCACGTCG 240

QY 572 TAATGCAACACCGGCTTACCCTCTGTGAAACAGCAGTCTTTCGGGTTCGAAGCAGAGTATC 631
 Db 241 TGCTGAGCACCGGCCACCTCTGCGAGCCAGCTGTCTCGGTGTGGAGGCCGAGTACC 300

QY 632 TCAGTCCGCTTTGATAAAAGTCAAGTGCAGAACAGAACTTCCCAAGAAAAAGAAATTC 691
 Db 301 TAAGTCCCTTTGATAAAAGCTCTGGAGCCCAACAGAGCCAGCATTTGGAGAAGAGTGGCGAAG 360

QY 692 A---AATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGTATGTTG 748
 Db 361 ACCCGAGGAGTGTGAGCTGTGATGTGTGGCGAGACATTTCCAGTGGCTTTTGTATGTTG 420

QY 749 AGATCCACATGAGAAACACACAAAGATTCTTTCACCTTACGGGTGTAAACATGTGCGGAAGAA 808
 Db 421 AGAGCCATGAGAGAGCATTAAGNCTCTTTCACGTATGGGTGCGATGTGCGGGAGGA 480

QY 809 GATTCAAGGAGCCTTGGTCTTTCTTAAAAATCATGCGGACRCATTAATGCGAAATCGGGG 868
 Db 481 GATTCAAGGAGCGGTGGTCTTCTGAAGAACACATGCGGACACACAATGGCAAGTCTGGCA 540

QY 869 CCAGAGCAAACTGCGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTCTGCC 928
 Db 541 CCAGAGCAAGCTTTCAGCAAGGCGATGGAG---AGTCCAGTCAACATCAATGAAGTGGTCC 597

QY 929 AGGTGACGCGCGGAGAGCATCTCTCTCTTGCAAAATCTGCATGTGTTGTGGCTTCC 988
 Db 598 AGCGCAGCCCTTGGGAGCATCTCCAGCCCTACAGATCTGCATGTCTGCGGCTTCC 657

QY 989 TATTTCCAAATAAAGAAAGTCTAATTGAGACCGCGCAAGGTGACACCAAAAAAATCTGTT 1048
 Db 658 TCTTCCCAATAAAGCAGAGCCCTCATTTGAGCAGCAAGGTTTCAACGCCAAAGAAATCTGTC 717

QY 1049 TCGGTACCGAGCGCGCAGACAGCATCTCCACAGGAGGATGCCGTCTCGAGGGAGG 1108
 Db 718 CCAGTGCCAGCAAGCTTGGCCCTGATGATCCACCGAGAGAAACCCACGTCCTCCGAGGAG 777

QY 1109 ACTTCTCTCAGTGTGTTCAACTTTGAGACCAAAAATCTCACCTCTGAAACGGGGGAAGACCTG 1168
 Db 778 AGTTGCTCAGTGTGTTGAACTTTGAGACCCAGATCAATCGCANGTAGTACAGTGAAGCCCA 837

QY 1169 TCAGATGATCCCTCAGCTCGATCGTTTACACCTTCCAGGCTTGGCAKCTGGGTAC 1226
 Db 838 TGACCTGATACCTCAGCTTGACCCGTTTACACACCTTACCGGCATGGCAGTTGGCTAC 895

RESULT 6

AW753609/c
 LOCUS RC1-CT0268-060100-013-h08 CT0268 Homo sapiens cDNA, mRNA sequence.
 DEFINITION
 ACCESSION AW753609
 VERSION AW753609.1 GI:7668541
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 569)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
TITLE
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
JOURNAL
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE
 20202663
PUBMED
 10737800
COMMENT
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC1&t2=RC1-CT0268-
 060100-013-h08&t3=2000-01-06&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 49
 High quality sequence stop: 67.
FEATURES
 source
 1..569
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="CT0268"
 /note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
 SmaI; A mini-library was made by cloning products derived
 from ORESTES PCR (U.S. Letters Patent application No.
 196,716 - Ludwig Institute for Cancer Research) profiles
 into the pUC 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."
ORIGIN
 Query Match 28.8%; Score 434.2; DB 2; Length 569;
 Best Local Similarity 89.5%; Pred. No. 1.3e-117;
 Matches 510; Conservative 1; Mismatches 55; Indels 4; Gaps 4;
 QY 430 TGTCAATGAAGGGCCNCCTGTGTCAATTCGAGCTACACAAAGAAAAAATGTCAATC 489
 DB 569 TGTCCATGAAGGGACCGCTGTGTTCATTCGAGCGACACAGAATAAGATGTC-ATC 511
 QY 490 CGAATCGAGGGGAATATGCCCTTGGATTGCATGTTCTGCA-CCGACGCTTCACACATTC 548
 DB 510 CCACCTCGAGGGATATATGCCCTTGGATGCATGTTCTGCATGCCAGACCTGCACACATTC 451
 QY 549 AGAAGACCTTAATAACATGTCTTAATGCAACACACCGGCTAGCC-TCTGTGAACACAGAG 607
 DB 450 AGAAGACCTGAATCCACATGTCTGATGCGACACACCGGCTACCCCTCTGTGAACACAGAG 391
 QY 608 TTCTTCGGGTGAAGCAGAGTATCTCAGTCCGCTTGTATAAAGTCAAGTGGCAACAGAAC 667
 DB 390 TTCTTCGGGTGAAGCAGAGTATCTCAGTCCGCTTGTATAAAGTCAAGTGGCAACAGAAC 331
 QY 668 CTCGCCAGAAAGAAATTCGAGGAAATGAATTTAGCTGTGAGGTATGTGGGCGACAT 727
 DB 330 CTCGCCAGAAAGAAATTCGAGGCAATGAATTTAGCTGTGAGGTATGTGGGCGACAT 271
 QY 728 TTAGAGTCCGCTTTTGTGATGTTGAGATCCACATGAGAACACACAAAGATCTTTCACCTACG 787
 DB 270 TTAGAGTCCGCTTTTGTGATGTTGAGATCCACATGAGAACACACAAAGATTCGCTACGACG 211
 QY 788 GGTGTAACATATGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAANAATCATCATCGCGA 847

Db 210 GGTGTGACATGTGCGGTAGAAGATTAAAGAGCCTTGGTTCTTAAATATCATCATCGGA 151
 QY 848 CRCATAATGGCAAAATCGGGGGCCAGAACTCAGCAAGCGCTTGGAGAGTAGTCCAG 907
 Db 150 CACATAATGGCAAAATCGGGGGCCAGAACTCAGCAAGCGCTTGGAGAGTAGTCCAG 91
 QY 908 CAACGATCAACAGAGTGTGTCAGGTCGACGCGCGGCGAGAGCATCTCTCTCC-TTGCAA 966
 Db 90 CCACGATCAACAGAGTGTGTCAGGTCGACGCGCGGCGAGAGCATCTCTCTCTCGTTACCA 31
 QY 967 ATCTGCATGTTTGTGGCTTCTCTATTCCA 996
 Db 30 ATCTGCATGTTTGTGGCTTCTCTATTCCA 1
RESULT 7
 BP366185
LOCUS
 DEFINITION
 sapiens cDNA clone TIR09337, mRNA sequence.
ACCESSION
 BP366185
VERSION
 BP366185.1 GI:52296435
KEYWORDS
 EST.
SOURCE
 Homo sapiens (human)
ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
 1 (bases 1 to 585)
 Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
 Mizushima-Sugano,J., Nakai,K. and Sugano,S.
 Sequence comparison of human and mouse genes reveals a homologous
 block structure in the promoter regions
 Genome Res. 14 (9), 1711-1718 (2004)
JOURNAL
 Contact: Yutaka Suzuki
COMMENT
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: yszukui@ims.u-tokyo.ac.jp.
FEATURES
 source
 1..585
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="TIR09337"
 /tissue_type="lung"
 /cell_type="fibroblast"
 /cell_lines="TIG"
 /dev_stage="fetal"
 /clone_lib="Sugano cDNA library, fetal lung fibroblast
 TIG"
ORIGIN
 Query Match 26.6%; Score 401; DB 5; Length 585;
 Best Local Similarity 93.8%; Pred. No. 1e-107;
 Matches 482; Conservative 0; Mismatches 26; Indels 6; Gaps 6;
 QY 2 CAGGTGTCTGGATGACTTCTTGTCAATTAACACACTCATTCAATGAGACAAGAGC 61
 Db 73 CAGGTGTCTGGATGACTTCTTGTCAATTAACACACTCATTCAATGAGACAAGAGA 132
 QY 62 ACTAATGCTTTGTGCTGATTTCATATTTGAATCGAGGCATTGGGAACCCCTGTATGCTTGT 121
 Db 133 ACTAATGCTTTGTGCTGATTTCATATTTGAATCGAGGCATTGGGAACCCCTGTATGCTTGT 192
 QY 122 TTGTGGAAGAACACAGTGACACCATCATGAGCTCTCTAAAGTTTGAAGAGTTAGAGG 181
 Db 193 TTGTGGAAGAACACAGTGACACCATCATGAGCTCTCTAAAGTTTGAAGAGTTAGAGG 252
 QY 182 ACTATACACTTTCTTTTGAACCTTTTATAATAATTTGCTCTGG-TTTTGAACCCAGG 240
 Db 253 ACTATACACTTTCTTTTGAACCTTTTATAATAATTTGCTCTGGTTTGTGAACCCAGG 312


```

Db      286 ACTATACACTTTCTTTTGAACCTTTTATATAAATAATTTGCTCTGGTTTGTGAAACCCAGG 345
Qy      241 ACTGTTAGAG-GGGTGAAGTACAGAGTCTTAC-AGTGGCCTTAATCCAATCCAGAAATTCG 298
Db      346 GCTGTTAGAGGGTGAAGTACAGAGTCTTACAGTGGCCTTATCCAACTCCAGAAATTCG 405
Qy      299 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCAACTCAATCC 358
Db      406 CCAACGGAACTTTGAGATTATATGCAATCGAAAGTGACAGGAAACATGCAACTCAATCC 465
Qy      359 CTCCTTAATGTACATGAT-GGCCAAGAGTGATGGCAGCTCTCTTGGCCAGTCCGATGGAG 417
Db      466 CTCCTTAATGTACATGATGGCCAGAGTGATGGCAGCTCTCTTGGCCAGTCCGATGGAG 525
Qy      418 ATGGA-GATGCTTGTCAATGAAGGGCCNCTGTG-TCAATTCGGAGCTACACAAA 473
Db      526 ATGGAAGATGCTTGTCAATGAAGGGCCGCTGTGTTCATTCGAGCTACACAAA 583

RESULT 10
LOCUS   CB457791
DEFINITION 636 bp mRNA linear EST 26-MAR-2003
ACCESSION 715904 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.
VERSION  CB457791
KEYWORDS  EST.
SOURCE    Bos taurus (cow)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
          Bovinae; Bos.
REFERENCE 1 (bases 1 to 636)
AUTHORS   Smith,T.P.L., Roberts,A.J., Echternkamp,S.E., Chitko-McKown,C.G.,
          Wray,J.E. and Keefe,J.W.
TITLE     A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL   Unpublished (2003)
COMMENT   Contact: Smith TPL
          USDA, ARS, US Meat Animal Research Center
          PO Box 166, Clay Center, NE 68933-0166, USA
          Tel: 402 762 4366
          Fax: 402 762 4390
          Email: smitht@mail.marc.usda.gov
          Single pass sequencing. Bases called with phred v0.020425.c and
          trimmed with the aid of the trim_alt option. Vector identified with
          cross match v0.990329.
          Plate: FQ8073 Row: A Column: 4
          Seq primer: GTAATACGACTCATTATAGG.
          Location/Qualifiers
            1..636
              /organism="Bos taurus"
              /mol_type="mRNA"
              /db_xref="taxon:9913"
              /tissue_type="pooled"
              /lab_host="DH10B"
              /clone_lib="MARC 6BOV"
              /note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
              Library made with RNA pooled from multiple tissues
              including liver, lung, hypothalamus, pituitary, and
              placenta/endometrium."

ORIGIN
Query Match 25.0%; Score 376.2; DB 6; Length 636;
Best Local Similarity 76.4%; Pred. No. 2.7e-100;
Matches 488; Conservative 1; Mismatches 144; Indels 6; Gaps 2;

Qy      496 GAGGGGAATATGCCCTTGATGATGATGTTTCGACGCCAGACTTCACACATTTCAGAAAC 555
Db      1 GAGGGGTGATGCCCTTGACATGATGTTCTGCACTCAGAGCTTCACGCGCTCCGAGGAC 60
Qy      556 CTTAATAACATGTTTATGCAACACCGGCTACCTCTGTGAACCAAGCATGTTCTTCGG 615
Db      61 CTCAGCAAGCAGCTCTCTCTGCAAGCAGCGGCCACCTCTGTGAGCCTGCGCTCTCGGC 120

```

```

Qy      616 GTTGAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAGAACAGAACTCTCCCAAG 675
Db      121 GTCGAGGCGGAGTATCTCAGCCCTCTGGATAAAGTCCAGGTCCGCGAGGAGCTCTCCCAAG 180
Qy      676 GAAAAGAAATTCGAAGAAAATGAA---TTTACTGTGAGGTATGTGGGAGACAGATTAGA 732
Db      181 GACAAAGAGCTGCAGGAGAACGAAGAGCTGAGCTGTGAAGTGTGCGGGCAGACGTTCCGC 240
Qy      733 GTCGCTTTTGTGTTGAGATCCACATGAGAACACACAAGAGATTCTTTTCACTTACGGGTCT 792
Db      241 GTGGCTTTTCGACCTCGAGATCCATGAGAAGAACACAAAGGACTCGTTCACTGACGGGTCT 300
Qy      793 AACATGTCGGGAAGAGATTCAAGGAGCTTGGTTTCTTAAATAATCACATGCGGACRCAT 852
Db      301 CACATGTGCGGCGCGGTTCGAAGGAGCTTGGTTTCTTGAAGAAATCACATGCGGACGCAC 360
Qy      853 AATGCAAAATCGGGGCGCAGAAAGCAAACTGCAAGCAAGGCTTGGAGAGTAGTCCAGCAAG 912
Db      361 ACGGCAAGTCCGGGGCGAAGAGCATGCTGCAAGCCGCGCTGGAG---AGCCCCGCGACC 417
Qy      913 ATCAACGAGGTGCTCCAGGTGCAAGCGCGCAGAGACATCTCTCTCTTGCATAATCTGC 972
Db      418 ATCAACGAGGTGCTCCAGAGCCCGCGCGAGGCGCTCTCTCGCGGTATAGATATGC 477
Qy      973 ATGTTTGTGCTCTCTTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCAC 1032
Db      478 ATGTTCTCGGCTTCTGTTTCCAAATAAAGAAAGTCTCATCGATCACCGAAGATGCAC 537
Qy      1033 ACCAAAAAATCTCTTTCGTTACAGAGCGCGCAGACAGACTCTCCACAGAGAGGATG 1092
Db      538 ACCAAAGAAAGTGTCTTCCGGCGCCCGCGCAGAGTCGAAGAGCCCAAGAGGAGCATG 597
Qy      1093 CGTCTCTCAGAGGAGGACTTCTCTGAGTTGTTCAACTTG 1131
Db      598 CTGTCCCGGGGCGGAGCTGCTGCAAGTCTCTGAATCTG 636

RESULT 11
LOCUS   CV409050/c
DEFINITION RCI-HT0268-090500-017-c10 HT0268 Homo sapiens cDNA, mRNA sequence.
ACCESSION CV409050
VERSION    CV409050.1 GI:52804553
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE   1 (bases 1 to 594)
AUTHORS     Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
TITLE       Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL     Proc.Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE     20202863
PUBMED      10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. http://www.ludwig.org.br.
            Location/Qualifiers
              1..594

```

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0268"
/notes="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match      23.8%; Score 359.2; DB 7; Length 594;
Best Local Similarity 84.9%; Pred. No. 3.1e-95;
Matches 485; Conservative 1; Mismatches 70; Indels 15; Gaps 7;

QY 423 GATGCTTTGTCAATCAAGGGCCNCTGTTG-TCAATTCGGAGCTACACAAAGAAAAA 481
DB 581 GATGCTTTGGCAGTGAAGGGACCGCTGTTTCCATTCGGAGCTACACAGGAAGAAAT 522
QY 482 TGTCAATCCGAATCAGAGGGGAATATGCCCTTGG-----ATTGCATGTTCTGCAGCCAGACC 537
DB 521 GGTATCCAAATCTGAGGGTGATATGCCCTTGGGATTTTCATGGTTGCTGCAGCCAGACC 462
QY 538 TTCACAT-----TCAGAAGACCTTAATAACATGCTTAAATGCAACACCGGC-CTACCC 592
DB 461 TTCACATATGTCAGAGAGACCTTGAAGATACATGCTTAAATGCAACACCGGCACCTACCC 402
QY 593 TCTGTGAACACGACGCTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTC 652
DB 401 TCTGTGAACACGACGCTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTC 342
QY 653 AAGT-GCGAACAGAACCTCCCAAGAAAGAAATTCGAAGAAATGAATTTAGCTGTGAG 711
DB 341 AAGTGGCCAAACAGAACCTCCCAAGAAAGAAATTCGAAGAAATGAATTTAGCTGTGAG 282
QY 712 GTATGTGGGCAGACATTTAGAGTCGCTTTTGNATGTTGATGATCCACATGAGAACACAAA 771
DB 281 GTATGTGTGACACATTTAGAGTCGCTTTTGNATGTTGATGATCCACATGAGAACACAAA 222
QY 772 GATTTCTTTCACCTTACGGGTGTAACTGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTT 831
DB 221 GAGTCTTTTCACTTACGGGTGTAACTGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTT 162
QY 832 AAAAATCATATCGGACACATATATGCGCAAAATCGGGGGCCAG--AAGCAAACTGCAGCAAG 889
DB 161 ATAAATCACAGCGGACACATAATGGCTATGCGGGGGCCAGATAGCTATCTGTCAAGCAAG 102
QY 890 GCTTGGAGATGATCCAGCAAGATCAAGAGGTC--GTCCAGGTGCACGGGGCCGAGAG 947
DB 101 GCGTGGAGATGATGTCACGACGATCAACAGGTTCTGTCCAGTGTGCACGGGGCCGAGAG 42
QY 948 CATCTCTCTCTCTGCAAAATCTGCATGGTT 978
DB 41 CATCTCTCTCTGACATGATCTGCATGTTT 11

RESULT 12
CB460210/c
LOCUS 720064 MARC 6BOV Bos taurus cDNA 3', mRNA linear EST 26-MAR-2003
DEFINITION CB460210
ACCESSION CB460210.1 GI:29266594
VERSION EST.
KEYWORDS Bos taurus (cow)
SOURCE Bos taurus
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
REFERENCE 1 (bases 1 to 662)
AUTHORS Smith,T.P.L., Roberts,A.J., Echternkamp,S.E., Chitko-McKown,C.G.,

Wray,J.E. and Keele,J.W.
A second set of bovine ESTs from pooled-tissue normalized libraries
Unpublished (2003)
Contact: Smith JPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smithemail.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross match v0.990329.
Plate: FOY8073 row: A column: 4
Seq primer: TAGAAGGCACAGTCGAGG.
FEATURES
source
1..662
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 6BOV"
/notes="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including liver, lung, hypothalamus, pituitary, and
placenta/endometrium."

ORIGIN
Query Match      22.8%; Score 343.8; DB 6; Length 662;
Best Local Similarity 75.3%; Pred. No. 1.3e-90;
Matches 466; Conservative 3; Mismatches 140; Indels 10; Gaps 3;

QY 703 AGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGATGTTGAGA-TCCACATGAG 761
DB 661 AGCTGTGAAGTGTGCGGGCAGACGTTTCGCGTGGCTTCGACGTGCAGATTCACATGAA 602
QY 762 AACACAAAGAATCTTTTCACTTACGGGTGTAAACATGTGCGGAAGAAGATTCAAGGAGCC 821
DB 601 GAACACAAAGACTCGTTTACGTACGGGTGTACATGTGCGGCGCGGCTTCAAGGAGCC 542
QY 822 TTGGTTTCTTAAATCATATGCGGACACATAATGGCAAAATCGGGGGCCAGAGCAAACT 881
DB 541 CTGGTTTCTGAAGAATCACATGCGGACGACACGCGCAAGTCGGGGGGCGAAGAGCAGGT 482
QY 882 GCAGCAAGCTTGGAGAGTAGTCCAGCAACCATCAACAGGTGCTCCAGGTGCACGCCGC 941
DB 481 GCAGCCCGCTTGGAG---AGCCCGCGACCATCAACAGGTGCTGCAGAGGCCCGCGGC 425
QY 942 CGAGACATCTCCTCTCCTTGCAAAATCTGCATGGTTTGTGGCTTCTTATTTCCAAATAA 1001
DB 424 CGAGGCGCTCTCTCGCGGTATAAGATATGCATGCTCTGCGCTTCTTGTGTTTCCAAATAA 365
QY 1002 AGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAAAATGCTTTCGTTACAGCAG 1061
DB 364 AGAAAGTCTCATCGATCACCGGAAGATGCACCAAAAGAAGCTGCTTCGGGCCCGCTGG 305
QY 1062 CGCGCAGACAGACTCTCCACAAAGGAGGAATGCCGCTCTCGAGGGAGGACTTCTCTGCAGTT 1121
DB 304 CCCGAGATCGAAGCCCCACAGGAAGGATGCTGTGCCGGGGCGGAGCTGCTGCAGTT 245
QY 1122 GTTCAACTTGAGACCAAAATCTCACCTGAAACGGGGGAAGAGCCTGTGCAGATGCATCCC 1181
DB 244 CTTGAATCTGAGACCCACGCTCCCGCGGAGATGCCAAGAAGCCGCGCAAGTGATACC 185
QY 1182 TCAGCTCGATCGTTTACCACCTTCCAGGCTTGGCAKCTGGCTACCAAGGAAWAGTTGC 1241
DB 184 TCAGCTGAGCCCGTTTACCACCTTACAGGCTTGGCAGCTGGCCACCAAGGGCAAGGTGC 125
QY 1242 CATTTGCCAAGAAGTGAAGG-----AATTGGGGCAAGAAGGAGCACCAGCAACACGACGA 1295
DB 124 TGTGTGCCGAGAGGTGAAGGAGCAGCAGCCCGGCCAGAGGGCAGACCGACATATGATGA 65
QY 1296 TTCGAGTTTCCGAGAGGAG 1314
```


QY 902 GTCCAGC-AACGATCAACGAGTCTCTCCAGGTGCACGGCGCGAGGC 948
Db 303 GTCCAGCAACGATCAACGAGTCTCTCCAGGTGCACGGCGCGAGGC 350

RESULT 15
BF237807
LOCUS BF237807 768 bp mRNA linear EST 14-NOV-2000
DEFINITION 601842326F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4079935 5',
mRNA sequence.

ACCESSION BF237807
VERSION BF237807.1 GI:11151713
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 768)

TITLE NIH-MGC <http://mgs.nci.nih.gov/>.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LCM939 row: c column: 08

High quality sequence stop: 464.

Location/Qualifiers

1..768

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4079935"

/tissue_type="leiomyosarcoma cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_46"

/note="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2:

EcoRI; cDNA made by oligo-dT priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GGCAGAG(G). Size-selected >500bp for average insert size

1.8kb. Library constructed by Ling Hong in the laboratory

of Gerald M. Rubin (University of California, Berkeley)

using ZAP-cDNA synthesis kit (Stratagene) and Superscript

II RT (Life Technologies). Note: this is a NIH_MGC

Library."

ORIGIN

Query Match 19.2%; Score 289.6; DB 2; Length 768;
Best Local Similarity 96.7%; Pred. No. 2e-74;
Matches 325; Conservative 2; Mismatches 6; Indels 3; Gaps 3;

QY 1175 GCATCCCTCAGTCGATCCGTTCCACACCTTCCAGGCTTGGCAKCTGGCTACCAAGGAA 1234
Db 1 GCATCCCTCAGTCGATCCGTTCCACACCTTCCAGGCTTGGCAKCTGGCTACCAAGGAA 60

QY 1235 WAGTTGCCAT-TTGCAGAGAGTGAAGGAATTGGGCAAGAGGAGCACCACCAACGAC 1293
Db 61 AAGTTGCCATTTTGCAGAGAGTGAAGGAATCGGGGCAAGAGGAGCACCACCAACGAC 120

QY 1294 GATTGAGTTCCGAGAAGAGCTTGGAGAAACAAATAAGAACCAATTGTGCAGGCTCTCG 1353
Db 121 GATTGAGTTCCGAGAAGAGCTTGGAGAAACAAATAAGGCGAGTTGTGCAGGCTCTCG 180

QY 1354 CA-AGAGAAAGAGAGTGCACACTCCACGGCGAAGCGCCCTCCGTGGAGCGGGATCC 1412
Db 181 CATAGAGAAGAGAGTGCACACTCCACGGCGAAGCGCCCTCCGTGGAGCGGGATCC 240

QY 1413 CAAG-TTACCCAGTAGCAAGGAGAGCCCACTCACTGCTCCAGTCCGCGAAAGCTTTCA 1471
Db 241 CAGATTACCCAGTAGCAAGGAGAGCCCACTCACTGCTCCAGTCCGCGAAAGCTTTCA 300
QY 1472 GAACCTACCAACAGCTGGTCTTTGCACCTCCAGGGTCC 1507
Db 301 GAACCTACCAACAGCTGGTCTTTGCACCTCCAGGGTCC 336

Search completed: June 29, 2005, 05:14:37
Job time : 3252.1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 14:57:12 ; Search time 7484.35 Seconds
(without alignments)
16865.305 Million cell updates/sec

Title: US-08-731-499-4
Perfect score: 2605
Sequence: 1 CAAAGCTCGAATTAACCTC.....CTGGCGTTACCAACTTAAT 2605

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sta.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2605	100.0	2605	6	AR070327	AR070327 Sequence
2	2605	100.0	2605	6	BD085728	BD085728 Genes fro
3	1158.4	44.5	106736	9	AL354993	AL354993 Human DNA
4	1092	41.9	1122	9	BC006282	BC006282 Homo sapi
5	1076.2	41.3	1243	6	AR339439	AR339439 Sequence
6	1073	41.2	1087	6	E50594	E50594 Human prote
7	1073	41.2	1087	6	BD095367	BD095367 Human pro
8	786.8	30.2	1112	10	BC029181	BC029181 Mus muscu
9	697.6	26.8	116397	9	AL157902	AL157902 Human DNA
10	644	24.7	134465	9	AC100793	AC100793 Homo sapi
11	644	24.7	187723	9	AC016889	AC016889 Homo sapi
12	531	20.4	531	6	E50584	E50584 Human prote
13	531	20.4	531	9	AB014763	AB014763 Homo sapi
14	476	18.3	495	6	AR412671	AR412671 Sequence
15	476	18.3	495	6	AX969505	AX969505 Sequence
16	476	18.3	495	6	BD108224	BD108224 EST and e
17	453	17.4	487	6	AR412670	AR412670 Sequence
18	453	17.4	487	6	AX969504	AX969504 Sequence
19	453	17.4	487	6	BD108223	BD108223 EST and e

20	444	17.0	494	6	AR414265	AR414265 Sequence
21	444	17.0	494	6	AX971099	AX971099 Sequence
22	444	17.0	494	6	BD109818	BD109818 EST and e
23	408.4	15.7	197101	2	AC133242	AC133242 Rattus no
24	389.6	15.0	417	6	AR424010	AR424010 Sequence
25	389.6	15.0	417	6	AX984704	AX984704 Sequence
26	389.6	15.0	417	6	BD119563	BD119563 EST and e
27	386.4	14.8	400	6	AR424014	AR424014 Sequence
28	386.4	14.8	400	6	AX984708	AX984708 Sequence
29	386.4	14.8	400	6	BD119567	BD119567 EST and e
30	384.6	14.8	195444	2	AC025424	AC025424 Mus muscu
31	384.6	14.8	250353	10	AL590969	AL590969 Mouse DNA
32	362	13.9	401	6	AR424009	AR424009 Sequence
33	362	13.9	401	6	AX984703	AX984703 Sequence
34	362	13.9	401	6	BD119562	BD119562 EST and e
35	343.2	13.2	588	6	CQ715946	CQ715946 Sequence
36	328.2	12.6	383	6	AR424011	AR424011 Sequence
37	328.2	12.6	383	6	AX984705	AX984705 Sequence
38	328.2	12.6	383	6	BD119564	BD119564 EST and e
39	319.6	12.3	894	5	BC079766	BC079766 Xenopus l
40	315	12.1	983	5	BX950269	BX950269 Gallus ga
41	312.6	12.0	817	5	CR762222	CR762222 Xenopus t
42	312.6	12.0	970	5	BX930330	BX930330 Gallus ga
43	312.6	12.0	971	5	BX929492	BX929492 Gallus ga
44	312.6	12.0	999	5	BX930730	BX930730 Gallus ga
45	312.6	12.0	1001	5	CR406053	CR406053 Gallus ga

ALIGNMENTS

RESULT 1
AR070327
LOCUS AR070327 2605 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 4 from patent US 5892010.
ACCESSION AR070327
VERSION AR070327.1 GI:7221215
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2605)
AUTHORS Gray,J., Collins,C., Hwang,S.-i., Godfrey,T., Kowbel,D. and Rommens,J.
TITLE Genes from the 20Q13 amplicon and their uses
JOURNAL Patent: US 5892010-A 4 06-APR-1999;
FEATURES Location/Qualifiers
source
1..2605
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN	Query Match	Best Local Similarity	Score	DB	Length
	100.0%;	100.0%;	2605;	2605;	2605;
	Matches 2605;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CAAGCTCGAATTAACCTCCTCACTAAAGGGAACAAAGCTGGAGCTCCACCGGCGGCGG 60			
Db	1	CAAGCTCGAATTAACCTCCTCACTAAAGGGAACAAAGCTGGAGCTCCACCGGCGGCGG 60			
Qy	61	CGCTCTAGAACTAGTGGATCCCGGGCTGAGGAATTCGGACGAGCTGGGCTACTAC 120			
Db	61	CGCTCTAGAACTAGTGGATCCCGGGCTGAGGAATTCGGACGAGCTGGGCTACTAC 120			
Qy	121	GATGCGGATGAGTTCGAGTGGCGGCTGGCAGTATCGCTTCCACCCCTCTTTACGTTACA 180			
Db	121	GATGCGGATGAGTTCGAGTGGCGGCTGGCAGTATCGCTTCCACCCCTCTTTACGTTACA 180			
Qy	181	ACCGAATGTGACACTCGGCAAGCAGCTGGCCGCTGGTGTCTGGTGTCTGTCTT 240			
Db	181	ACCGAATGTGACACTCGGCAAGCAGCTGGCCGCTGGTGTCTGGTGTCTGTCTT 240			
Qy	241	CTGCCGCTGCACAAACAGTCCAGATGACGGTGTGGAAGCTCAGGAGAGCCGCTCTT 300			

Db 241 CTGCCGCTGCACAAACAGTCCAGCATGACGGTGATGAAGCTCAGGAGAGCCCGCTCTT 300
QY 301 CAAACAGCTCAAGCTACAGCGAAAGCTTCCTGTGAGTGCATCCAGATTGTTATAGAGGA 360
Db 301 CAAACAGCTCAAGCTACAGCGAAAGCTTCCTGTGAGTGCATCCAGATTGTTATAGAGGA 360
QY 361 ACTGAGGAAGAAAGGAAACCTCAGTGGTTGGATAAGAGCAAGTCACAGCTTCCTCATCAT 420
Db 361 ACTGAGGAAGAAAGGAAACCTCAGTGGTTGGATAAGAGCAAGTCACAGCTTCCTCATCAT 420
QY 421 GTGGCGGAGGCCAGAAAGATGGGGAAACTCATCTATCAGTGGGTTTCCAGGAGTGGCCA 480
Db 421 GTGGCGGAGGCCAGAAAGATGGGGAAACTCATCTATCAGTGGGTTTCCAGGAGTGGCCA 480
QY 481 GAAACAACCTCCGTCTTTACCTGTATGAACCTGACTAATGGGGAAGACACAGAGGATGAGGA 540
Db 481 GAAACAACCTCCGTCTTTACCTGTATGAACCTGACTAATGGGGAAGACACAGAGGATGAGGA 540
QY 541 GTTCCACGGCTGGATGAAGCCACTCTACTCGGGCTCTGCAGGCCCTACAGCGAGGACA 600
Db 541 GTTCCACGGCTGGATGAAGCCACTCTACTCGGGCTCTGCAGGCCCTACAGCGAGGACA 600
QY 601 CAAGGCCGAGATCATCACTGTGAGCGATGGCCGAGGCGTCAAGTTCTTCTAGCAGGGACC 660
Db 601 CAAGGCCGAGATCATCACTGTGAGCGATGGCCGAGGCGTCAAGTTCTTCTAGCAGGGACC 660
QY 661 TGTCTCCCTTTACTTCTTACCTCCACCTTTCAGGGCTTCAAAGGAGACAGACCCAG 720
Db 661 TGTCTCCCTTTACTTCTTACCTCCACCTTTCAGGGCTTCAAAGGAGACAGACCCAG 720
QY 721 TGTCCCCCAAGACATGGGATCTGTGACTCCACAGACTCAAAAGGACTCCAGTCCCTGAAGG 780
Db 721 TGTCCCCCAAGACATGGGATCTGTGACTCCACAGACTCAAAAGGACTCCAGTCCCTGAAGG 780
QY 781 CTGGGACCTGGGGATGGGTTTCTCACACCCCATATGTCTGCTCCCTTGGATAGGGTGAGGC 840
Db 781 CTGGGACCTGGGGATGGGTTTCTCACACCCCATATGTCTGCTCCCTTGGATAGGGTGAGGC 840
QY 841 TGAAGCACAGGGAGAAATATGTGGTCTTCCTCGCCCTACCTCTTCCCATCTCCTAGAC 900
Db 841 TGAAGCACAGGGAGAAATATGTGGTCTTCCTCGCCCTACCTCTTCCCATCTCCTAGAC 900
QY 901 TGTCTCTGAGCCAGGGTCTGTAACCTGACACTTTATATGTTTACACATGTAAGTACA 960
Db 901 TGTCTCTGAGCCAGGGTCTGTAACCTGACACTTTATATGTTTACACATGTAAGTACA 960
QY 961 TACACACATGCGCCTGCAGCACATGCTTCTGTCTCTCTCTCCACCCCTTTTAGCTGC 1020
Db 961 TACACACATGCGCCTGCAGCACATGCTTCTGTCTCTCTCTCTCTCCACCCCTTTTAGCTGC 1020
QY 1021 TGTTCGCTCCCTCTCAGCGTGGTCTGTGATCCTTCTAGGGGATGGGGAGCCCTGCG 1080
Db 1021 TGTTCGCTCCCTCTCAGCGTGGTCTGTGATCCTTCTAGGGGATGGGGAGCCCTGCG 1080
QY 1081 TGCAGGACGCTTCAGGCAATATGAAGATAGGAGCCACAGGGCTGCGAGTGAGAGGT 1140
Db 1081 TGCAGGACGCTTCAGGCAATATGAAGATAGGAGCCACAGGGCTGCGAGTGAGAGGT 1140
QY 1141 GTGGCCCCACACCGGATTTATGATATATAAATCTCAACTCCCAAAAAAAGGATGAGGT 1200
Db 1141 GTGGCCCCACACCGGATTTATGATATATAAATCTCAACTCCCAAAAAAAGGATGAGGT 1200
QY 1201 CTGAGACTAGTCTCTCTCTCGAGAACTAGTCTCGAGTTTCTTCTTTTCTTTTCTTTT 1260
Db 1201 CTGAGACTAGTCTCTCTCTCGAGAACTAGTCTCGAGTTTCTTCTTTTCTTTTCTTTT 1260
QY 1261 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1320
Db 1261 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1320
QY 1321 ACTTTTCTCTACTTAATCTACTTTCCAGTCTCAGAGCCCAAGAGGAAAAAAGAC 1380

Db 1321 ACTTTTCTCTACTTAATCTACTTTTCCAGTCTCAGAGCCCAAGAGGAAAAAAGAC 1380
QY 1381 CATGAATCTTCTCTCCAGATTAAAGTACACACTTTTGGAAAAACAGATTGGAAACCTTT 1440
Db 1381 CATGAATCTTCTCTCCAGATTAAAGTACACACTTTTGGAAAAACAGATTGGAAACCTTT 1440
QY 1441 CTGAAAAAGTTGACTGAAACTCCAAACCAACATGCCATATTTGTTGATGTTGCTCATGAA 1500
Db 1441 CTGAAAAAGTTGACTGAAACTCCAAACCAACATGCCATATTTGTTGATGTTGCTCATGAA 1500
QY 1501 AATTGTTAAAAACCTGTTCTAGATAAAGAACAGTCTCAAGTTTGTAGAGCTTACAT 1560
Db 1501 AATTGTTAAAAACCTGTTCTAGATAAAGAACAGTCTCAAGTTTGTAGAGCTTACAT 1560
QY 1561 AGTACAAGGGTCCCTATGATGATCTTCTGTAGAGCAAAATAATGTAATTTTTTCAGTT 1620
Db 1561 AGTACAAGGGTCCCTATGATGATCTTCTGTAGAGCAAAATAATGTAATTTTTTCAGTT 1620
QY 1621 TCTGTTTTATAACTCTCTCGATCTCAGAGTTGACTGATTAAAAACACCTACTCATGCAACA 1680
Db 1621 TCTGTTTTATAACTCTCTCGATCTCAGAGTTGACTGATTAAAAACACCTACTCATGCAACA 1680
QY 1681 GAGATAAAGCACTCATATTTTTTATAAATATATAGGACCAAACTATTTTGGAAATCTTAT 1740
Db 1681 GAGATAAAGCACTCATATTTTTTATAAATATATAGGACCAAACTATTTTGGAAATCTTAT 1740
QY 1741 CTATTGGAGACATATGCTGGACTAAAGCAATAATTTTATTTCTCAATGTCGTGC 1800
Db 1741 CTATTGGAGACATATGCTGGACTAAAGCAATAATTTTATTTCTCAATGTCGTGC 1800
QY 1801 TAACTCTCAATGACTTAGAATGCTTTGCTATATTTTTTGCCTCTATGCCTCAACACACTGCG 1860
Db 1801 TAACTCTCAATGACTTAGAATGCTTTGCTATATTTTTTGCCTCTATGCCTCAACACACTGCG 1860
QY 1861 TTTCTTTTAGCTCTTGAACAGCCAAACTGCTTCTGCTCAGGACAGATTTTGGGA 1920
Db 1861 TTTCTTTTAGCTCTTGAACAGCCAAACTGCTTCTGCTCAGGACAGATTTTGGGA 1920
QY 1921 CTTCTCTTAAGANTCTAATTTCTTAATTTCTTAATCTGGGTAACTTAGTTTATCCAACA 1980
Db 1921 CTTCTCTTAAGANTCTAATTTCTTAATTTCTTAATCTGGGTAACTTAGTTTATCCAACA 1980
QY 1981 CTTAGATCTCTGCCGTAATAAACTCTTTATAGAGCCCTGTCATGACACTGCTCTCTTC 2040
Db 1981 CTTAGATCTCTGCCGTAATAAACTCTTTATAGAGCCCTGTCATGACACTGCTCTCTTC 2040
QY 2041 TCCAACTACTCACCAGCACACATGTAGACTAGATTAGAACTCTCTGTTTTCTTTTCA 2100
Db 2041 TCCAACTACTCACCAGCACACATGTAGACTAGATTAGAACTCTCTGTTTTCTTTTCA 2100
QY 2101 TACTTTTCTCTATCATGCTTCCCTCCATTAATAATTTTTTATATGTTGTAATGTCG 2160
Db 2101 TACTTTTCTCTATCATGCTTCCCTCCATTAATAATTTTTTATGTTGTAATGTCG 2160
QY 2161 CCCCAAGTCAGTTTCTCTCACTAAACTATAAACTCCGTAAGCTGGGATCTTCCAAATTT 2220
Db 2161 CCCCAAGTCAGTTTCTCTCACTAAACTATAAACTCCGTAAGCTGGGATCTTCCAAATTT 2220
QY 2221 GATCACCACTTAGTACAGTAGGAAACACAGTAAAGATTCAATTTGGTATTTTGTGGAATGAT 2280
Db 2221 GATCACCACTTAGTACAGTAGGAAACACAGTAAAGATTCAATTTGGTATTTTGTGGAATGAT 2280
QY 2281 GAAATGAATTTTGTCTAGTAAAGTCTGGGGAAACCCAGGTGAGAGAGCCCTAGAAAGCA 2340
Db 2281 GAAATGAATTTTGTCTAGTAAAGTCTGGGGAAACCCAGGTGAGAGAGCCCTAGAAAGCA 2340
QY 2341 GGTTCGAATCCAAGGCTAGATAGACTTAGTGTCTTACTCAAGAAGGGTAGCCTGAAATAAA 2400
Db 2341 GGTTCGAATCCAAGGCTAGATAGACTTAGTGTCTTACTCAAGAAGGGTAGCCTGAAATAAA 2400
QY 2401 GGTTCGAATTTAGTCAAGAAATAGTCAAGACATGGGCAAGACAGAGTGTCTGTCGCC 2460
Db 2401 GGTTCGAATTTAGTCAAGAAATAGTCAAGACATGGGCAAGACAGAGTGTCTGTCGCC 2460

QY 2461 GAATTCGATATCAAGCTTATTCGATACCGTCGACCTCGAGGGGGGGCCCCGGTACCCCAATTC 2520
DB 2461 GAATTCGATATCAAGCTTATTCGATACCGTCGACCTCGAGGGGGGGCCCCGGTACCCCAATTC 2520
QY 2521 GCCTATAGTAGTCTGATTTACAAATTCACCTGGCGGTCGTTTACAAAGCTCGTGACTGGGA 2580
DB 2521 GCCTATAGTAGTCTGATTTACAAATTCACCTGGCGGTCGTTTACAAAGCTCGTGACTGGGA 2580
QY 2581 AAACCTGGCGTTTACCCTTAAT 2605
DB 2581 AAACCTGGCGTTTACCCTTAAT 2605

RESULT 2
BD085728
LOCUS BD085728 2605 bp DNA linear PAT 27-AUG-2002
DEFINITION Genes from the 20q13 amplicon and their uses.
ACCESSION BD085728
VERSION BD085728.1 GI:22631338
KEYWORDS JP 2001524802-A/4.
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences: artificial sequences.
REFERENCE 1 (bases 1 to 2605)
AUTHORS Gray,J.W., Collins,C.C., Hwang,S.I., Godfrey,T., Kowbel,D. and Rommens,J.
TITLE Genes from the 20q13 amplicon and their uses
JOURNAL Patent: JP 2001524802-A 4 04-DEC-2001;
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
OS Artificial Sequence
COMMENT PN JP 2001524802-A/4
PD 04-DEC-2001
PF 15-JUL-1997 JP 1998056264
PR 15-JUL-1996 US 08/680395,16-OCT-1996 US 08/731499 PR
17-JAN-1997 US 08/785532
PI JOE W GRAY, COLIN CONRAD COLLINS, SOO IN HWANG, TONY GODFREY, PI DAVID KOWBEL,
PI JOHANNA ROMMENS
CC C12N15/11,C12Q1/68,A61K48/00
PC Description of Artificial Sequence:cc43
FH Key Location/Qualifiers
FT source 1..2605
FT /organism='Artificial Sequence'.
FEATURES
source 1..2605
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 100.0%; Score 2605; DB 6; Length 2605;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CAAGCTCGAAATTAACCTCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGGTGGCGG 60
DB 1 CAAGCTCGAAATTAACCTCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGGTGGCGG 60
QY 61 CGCTCTAGAACTAGTGGATCCCGGGCTGCGGAAATTCGGACGAGCTGGGCTACTAC 120
DB 61 CGCTCTAGAACTAGTGGATCCCGGGCTGCGGAAATTCGGACGAGCTGGGCTACTAC 120
QY 121 GATGGGATGAGTTTCGAGTGGCGGCGGATATCGCTTCCACCCCTCTTTACGTTTACA 180
DB 121 GATGGGATGAGTTTCGAGTGGCGGCGGATATCGCTTCCACCCCTCTTTACGTTTACA 180
QY 181 ACCGAAATGGACACTCGGACAGAGCTGGCGGCTGGTGTCTGTCTGTCTGTCTGTCTGT 240
DB 181 ACCGAAATGGACACTCGGACAGAGCTGGCGGCTGGTGTCTGTCTGTCTGTCTGTCTGT 240
QY 241 CTGGCCCTTGCAAAACAGTCCAGCATGACGGTATGGAAGCTCAGGAGAGCCCGCTCTT 300

DB 241 CTGGCCCTTGCAAAACAGTCCAGCATGACGGTATGGAAGCTCAGGAGAGCCCGCTCTT 300
QY 301 CAACAAACGTCAAGCTACAGCGAAAGCTTCTCTGTGGAGTCGATCCAGATTTGTTATGAGGA 360
DB 301 CAACAAACGTCAAGCTACAGCGAAAGCTTCTCTGTGGAGTCGATCCAGATTTGTTATGAGGA 360
QY 361 ACTGAGGAAGAAAGGAACCTCGAGTGGTTGATTAAGACAAGTCCAGCTTCTGTATCAT 420
DB 361 ACTGAGGAAGAAAGGAACCTCGAGTGGTTGATTAAGACAAGTCCAGCTTCTGTATCAT 420
QY 421 GTGGCGGAGGCCAGAAAGTGGGGAACCTCATCTATCAGTGGGTTTCAGGAGTGGCCA 480
DB 421 GTGGCGGAGGCCAGAAAGTGGGGAACCTCATCTATCAGTGGGTTTCAGGAGTGGCCA 480
QY 481 GAACAACTCCGTCTTTTACCTCTGATGAACCTGATTAATGGGGAAGACACAGAGGATGAGGA 540
DB 481 GAACAACTCCGTCTTTTACCTCTGATGAACCTGATTAATGGGGAAGACACAGAGGATGAGGA 540
QY 541 GTTCCACGGGCTGGATGAAGCCACTCTACTCGGGCTCTGCAGGCCCTTACAGCAGGAGCA 600
DB 541 GTTCCACGGGCTGGATGAAGCCACTCTACTCGGGCTCTGCAGGCCCTTACAGCAGGAGCA 600
QY 601 CAAGCCGAGATCATCACTGTCTAGCGATGGCGAGCGTCAAGTTCTTCTAGCAGGAGACC 660
DB 601 CAAGCCGAGATCATCACTGTCTAGCGATGGCGAGCGTCAAGTTCTTCTAGCAGGAGACC 660
QY 661 TGTCTCCCTTTACTTTTACCTCCACCTTTCCAGGGCTTTCAAAGGAGACAGACCCAG 720
DB 661 TGTCTCCCTTTACTTTTACCTCCACCTTTCCAGGGCTTTCAAAGGAGACAGACCCAG 720
QY 721 TGTCCCCCAAAGACTGGATCTGTGACTCACACAGACTCAAAAGGACTCCAGCTCTGAAGG 780
DB 721 TGTCCCCCAAAGACTGGATCTGTGACTCACACAGACTCAAAAGGACTCCAGCTCTGAAGG 780
QY 781 CTGGGACCTGGGGATGGGTTTCTCAACCCCATATGTCTGTCTCTTGGATAGGTCGAGGC 840
DB 781 CTGGGACCTGGGGATGGGTTTCTCAACCCCATATGTCTGTCTCTTGGATAGGTCGAGGC 840
QY 841 TGAAGCACCCAGGGAGAAAATATGTCTCTCGCCCTACCTCTCTTCCCATCTCTAGAC 900
DB 841 TGAAGCACCCAGGGAGAAAATATGTCTCTCTCGCCCTACCTCTCTTCCCATCTCTAGAC 900
QY 901 TGTCTTTGAGCCAGGCTGTGTAACCTGACACTTATATGTGTTCACATGTAAGTACA 960
DB 901 TGTCTTTGAGCCAGGCTGTGTAACCTGACACTTATATGTGTTCACATGTAAGTACA 960
QY 961 TACACACATGCGCCCTGCAGCACATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
DB 961 TACACACATGCGCCCTGCAGCACATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
QY 1021 TGTTCCTCCCTTCTCAGGCTGGTGTGATCTTCTCTAGGGGATGGGGGAAGCCCTGGC 1080
DB 1021 TGTTCCTCCCTTCTCAGGCTGGTGTGATCTTCTCTAGGGGATGGGGGAAGCCCTGGC 1080
QY 1081 TGCAGGAGCCCTTCAGGCAATATGAAGATAGGAGGCCACCGGGCTGGCAGTGAGAGT 1140
DB 1081 TGCAGGAGCCCTTCAGGCAATATGAAGATAGGAGGCCACCGGGCTGGCAGTGAGAGT 1140
QY 1141 GTGGCCCCCAGCCGATTTATGATATTAATCTCACTCCCAAAAAAAGGAAAAA 1200
DB 1141 GTGGCCCCCAGCCGATTTATGATATTAATCTCACTCCCAAAAAAAGGAAAAA 1200
QY 1201 CTGAGACTAGTTCTCTCTCTCGAGAACTAGTCTCGAGTCTTCTTCTTCTTCTTCTTCT 1260
DB 1201 CTGAGACTAGTTCTCTCTCTCGAGAACTAGTCTCGAGTCTTCTTCTTCTTCTTCTTCT 1260
QY 1261 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1320
DB 1261 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1320
QY 1321 ACTTTTCTTCTACTTAACTACTTTTCTCTCAGAGCCCGGAGGAAAAAAGAAC 1380
DB 1321 ACTTTTCTTCTACTTAACTACTTTTCTCTCAGAGCCCGGAGGAAAAAAGAAC 1380

1381 CATGAATCTTCTCTCCAGATTAAAGTACACACTTTTGGAAAAAGATTTGGAAAAACCTTTT 1440
Db CATGAATCTTCTCTCCAGATTAAAGTACACACTTTTGGAAAAAGATTTGGAAAAACCTTTT 1440
1441 CTGAAAAAAGTTGACTGAAACTCCAAACCAACATGCCATATTTGTTGATGTTGCTCATGAA 1500
Db CTGAAAAAAGTTGACTGAAACTCCAAACCAACATGCCATATTTGTTGATGTTGCTCATGAA 1500
1501 AATTGTTAAAAACCTGTTCTAGATAAAGAACAGTCTCAAGTTTGTATCAGCCTACACAT 1560
Db AATTGTTAAAAACCTGTTCTAGATAAAGAACAGTCTCAAGTTTGTATCAGCCTACACAT 1560
1561 AGTACAAGGTCCTCTATGATGATCTCTGCTAGACGAAATAAATGTAATTTTTTCAGTT 1620
Db AGTACAAGGTCCTCTATGATGATCTCTGCTAGACGAAATAAATGTAATTTTTTCAGTT 1620
1621 TCTGTTTTATAACTCTCTCGATCTCAGAGTTGACTGATTAAACCACTACTCATGCAACA 1680
Db TCTGTTTTATAACTCTCTCGATCTCAGAGTTGACTGATTAAACCACTACTCATGCAACA 1680
1681 GAGATAAAGCACTCATATTTTATTAATATATGACCAAACTATTTTGGAAATCTTAT 1740
Db GAGATAAAGCACTCATATTTTATTAATATATGACCAAACTATTTTGGAAATCTTAT 1740
1741 CTATTGGAGACAAATATGCTGACCTAAAGCAATAATTTTATCTCAATGCTGTGTC 1800
Db CTATTGGAGACAAATATGCTGACCTAAAGCAATAATTTTATCTCAATGCTGTGTC 1800
1801 TAACTCAATGACTTTAGAAATGCTTTGCTATATTTTGCCTCTATGCTCAACACACTGCG 1860
Db TAACTCAATGACTTTAGAAATGCTTTGCTATATTTTGCCTCTATGCTCAACACACTGCG 1860
1861 TTTCTTTTGTCTTTGAACAGCAAACTGCTTCTGCTCAGGACCAAGATATTTTGGGA 1920
Db TTTCTTTTGTCTTTGAACAGCAAACTGCTTCTGCTCAGGACCAAGATATTTTGGGA 1920
1921 CTTCTCTTAAGAAATCTATTTCTTAATCTTTATCTGGTAACTAGTTTATCCAACA 1980
Db CTTCTCTTAAGAAATCTATTTCTTAATCTTTATCTGGTAACTAGTTTATCCAACA 1980
1981 CTTGAGATCTGCGCGTAAAAACTCTTTATAGAAAGCCTGTATGACACTGCTCTCTTC 2040
Db CTTGAGATCTGCGCGTAAAAACTCTTTATAGAAAGCCTGTATGACACTGCTCTCTTC 2040
2041 TCCACATACTCACAGACACACATGACTAGATAGAACTCTCTGTTTCTTTTCA 2100
Db TCCACATACTCACAGACACACATGACTAGATAGAACTCTCTGTTTCTTTTCA 2100
2101 TACTTTTCTCTATCATGCTTCTCCATATATATATTTTATATGTTGTAATGCTG 2160
Db TACTTTTCTCTATCATGCTTCTCCATATATATATTTTATATGTTGTAATGCTG 2160
2161 CCCCAAGTCAGTTTCTCTCACTAAACTATAAACTCCGTAAGCTGGATCCCTTCCAAATTT 2220
Db CCCCAAGTCAGTTTCTCTCACTAAACTATAAACTCCGTAAGCTGGATCCCTTCCAAATTT 2220
2221 GATCACCTTAGTACAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2280
Db GATCACCTTAGTACAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 2280
2281 GAATCAATTTGTTTCTAGTAAAGTCTGGGGAACCCAGGTGAGAAAGCCTAGAAAGCA 2340
Db GAATCAATTTGTTTCTAGTAAAGTCTGGGGAACCCAGGTGAGAAAGCCTAGAAAGCA 2340
2341 GGTGGAATCCAGGCTAGATAGACTTAGTGTGTTACTCAAGAAAGGCTAGCTGCTGCTGCC 2400
Db GGTGGAATCCAGGCTAGATAGACTTAGTGTGTTACTCAAGAAAGGCTAGCTGCTGCTGCC 2400
2401 GGTTCAAATATAGTCAAGAAATAGTCAAGACATGGGCAAGACAGAGTCTGCTGCTGCC 2460
Db GGTTCAAATATAGTCAAGAAATAGTCAAGACATGGGCAAGACAGAGTCTGCTGCTGCC 2460

2461 GAAATTCGATATCAAGCTTATCGATACCGTCCGAGGGGGGGGGGGGGGGGGGGGGGGGGGG 2520
Db GAAATTCGATATCAAGCTTATCGATACCGTCCGAGGGGGGGGGGGGGGGGGGGGGGGGGGG 2520
2521 GCCCTATAGTGAAGTCTGATTAACAATTCATCTGCGCTGCTGTTTACAACTGCTGACTGGGA 2580
Db GCCCTATAGTGAAGTCTGATTAACAATTCATCTGCGCTGCTGTTTACAACTGCTGACTGGGA 2580
2581 AAACCTGGCGTTACCCCAACTTAAT 2605
Db AAACCTGGCGTTACCCCAACTTAAT 2605

AL354993 106736 bp DNA linear PRI 20-JUN-2001
Human DNA sequence from clone RP5-823G15 on chromosome 20q13.2-13.3
Contains a peptidylprolyl isomerase A (cyclophilin A) (PPIA)
pseudogene, the gene for OVC10-2, part of a gene encoding a
putative zinc finger protein ZNF218, ESTs, STSS and GSSs, complete
sequence.
AL354993 GI:11121037
AL354993.24 GI:11121037
HTG; cyclophilin; OVC10-2; peptidoprolyl isomerase; PPIA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 106736)
AUTHORS Frankland, J.
TITLE Direct Submission
JOURNAL Submitted (15-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Nov 8, 2000 this sequence version replaced gi:11096473.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WormPEP; Information
on the WormPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 20, constructed by the Sanger Centre Chromosome 20
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr20
IMPORTANT: This sequence is not the entire insert of clone
RP5-823G15 It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true left end of clone RP4-724E16 is at 106637 in this
sequence. The true right end of clone RP4-669H2 is at 100 in this
sequence. This sequence was finished as follows unless otherwise
noted: all regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., phred
quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.
RP5-823G15 is from the library RPCI-5 constructed by the group of
Pister de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2.
Location/Qualifiers
1. 106736
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="q13.2-13.3"

QY 1279 TGGCTTTAAGGATTTATTTATTTGTTTCCCTTTACAGTGTCCACTTTTCTCTACTTAATA 1338
Db 45960 TGGCTTTAAGGATTTATTTA--GTTTCCCTTTACAGTGTCCACTTTTCTCTACTTAATA 46017
QY 1339 CTACTTTTCAGTCTCAGAACCCAGAGGCGAAAAAAGACACATGATCTTCTCTCTCC 1398
Db 46018 CTACTTTTCAGTCTCAGAACCCAGAGGCGAAAAAAGACACATGATCTTCTCTCTCC 46077
QY 1399 AGATTAAAGTACACACTTTTGAAAAACAGATTGGAAACCTTTTCTGAAAAAGTTGACTGA 1458
Db 46078 AGATTAAAGTACACACTTTTGAAAAACAGATTGGAAACCTTTTCTGAAAAAGTTGACTGA 46137
QY 1459 RACTCCAAACCAACATGCGATTTCTGATGTGCTCATGAATAATGTTTAAAAAAGCTGTT 1518
Db 46138 RACTCCAAACCAACATGCGATTTCTGATGTGCTCATGAATAATGTTTAAAAAAGCTGTT 46197
QY 1519 CTAGATAAAGAACAGTCTCAAGCTTTTGTACAGCCTTACACATAGTACAAGGTCCTCCCTAT 1578
Db 46198 CTAGATAAAGAACAGTCTCAAGCTTTTGTACAGCCTTACACATAGTACAAGGTCCTCCCTAT 46257
QY 1579 GATGATCTCTGTAGGACGAAATAATGTAATTTTTCAGTTTCTGGTTTATAACTCTCT 1638
Db 46258 GATGATCTCTGTAGGACGAAATAATGTAATTTTTCAGTTTCTGGTTTATAACTCTCT 46317
QY 1639 CGATCTCAGATTGACTGATTAACACACTACTCATGCAACAGAGAAATAAGACCTCAT 1698
Db 46318 CGATCTCAGATTGACTGATTAACACACTACTCATGCAACAGAGAAATAAGACCTCAT 46377
QY 1699 TTTTATAAATATATATGACCAAACTATTTTGGAAATCTTATCTATGAGACACAAATAT 1758
Db 46378 TTTTATAAATATATGACCAAACTATTTTGGAAATCTTATCTATGAGACACAAATAT 46437
QY 1759 GCTGACTAAAGCAATAATATTTATTTCTCAATGTCTGTGTAAACCTCAATGACTTGA 1818
Db 46438 GCTGACTAAAGCAATAATATTTATTTCTCAATGTCTGTGTAAACCTCAATGACTTGA 46497
QY 1819 ATGCTTTGCTATATTTTGCCTCTATGCTCAACACACCTGGCTTCTTTAGCTCTTCAA 1878
Db 46498 ATGCTTTGCTATATTTTGCCTCTATGCTCAACACACCTGGCTTCTTTAGCTCTTCAA 46557
QY 1879 CAAGCCAAACTGCTTCCCTCGCTCAGGACACAGATATTTTGGGACTTCTCTTAAAGATTCTA 1938
Db 46558 CAAGCCAAACTGCTTCCCTCGCTCAGGACACAGATATTTTGGGACTTCTCTTAAAGATTCTA 46617
QY 1939 TTTCTCTTAATCTTATCTGGGTAACTTAGTTTTATCCAACTTCAGATCCGCGTAA 1998
Db 46618 TTTCTCTTAATCTTATCTGGGTAACTTAGTTTTATCCAACTTCAGATCCGCGTAA 46677
QY 1999 AAACCTCTTATAGAGCCTGTCATGACACCTCTCTCTTCCAACTATCTCACCAGC 2058
Db 46678 AAACCTCTTATAGAGCCTGTCATGACACCTCTCTCTTCCAACTATCTCACCAGC 46737
QY 2059 ACACATGTAGACTAGTAACCTCTGTTTCTTTTCTTCTTCTATCTATCATGC 2118
Db 46738 ACACATGTAGACTAGTAACCTCTGTTTCTTTTCTTCTTCTATCTATCATGC 46797
QY 2119 TTCCCTCCATTAATATTTATATGTTGTTGAATGTCTGCCCAAGTCAAGTTTCTCTC 2178
Db 46798 TTCCCTCCATTAATATTTATATGTTGTTGAATGTCTGCCCAAGTCAAGTTTCTCTC 46857
QY 2179 ACTAAACTATAAATCCGTAAGCTGGGATCCTTCCAAATTTTGTACACCACTTAGTACAG 2238
Db 46858 ACTAAACTATAAATCCGTAAGCTGGGATCCTTCCAAATTTTGTACACCACTTAGTACAG 46917
QY 2239 TAGGAACACAGTAAAGATTCAATTTGGTATTTTGTGGAATGAATGAATTTGTTTGTCTA 2298
Db 46918 TAGGAACACAGTAAAGATTCAATTTGTTGGAATGAATGAATTTGTTTGTCTA 46977
QY 2299 GTAAAGTCTGGGGGAAACCCAGGTGAGAGAGCCTTGAAGAAGCGAGTCTGAATCCAAAGCTAG 2358
Db 46978 GTAAAGTCTGGGGGAAACCCAGGTGAGAGAGCCTTGAAGAAGCGAGTCTGAATCCAAAGCTAG 47037

QY 2359 ATAGACTTAGTGTACTCAAGAAAGGCTAGCTGAAAAATAAGGTTCAAATTTATAGTCAA 2418
Db 47038 ATAGACTTAGTGTACTCAAGAAAGGCTAGCTGAAAAATAAGGTTCAAATTTATAGTCAA 47097
QY 2419 GAATAGTCAAGACATGCGGCAAGACAAAGAGTGTGTG 2452
Db 47098 GAATAGTCAAGACATGCGGCAAGACAAAGAGTGTGTG 47131

RESULT 4
BC006282
LOCUS
DEFINITION
MGC:10540 IMAGE:4025884, complete cds.
ACCESSION
BC006282
VERSION
BC006282.2 GI:34785072
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE
1 (bases 1 to 1122)
AUTHORS
Straussberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Wax, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schaefer, T.E., Brownstein, M.J., Udutin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Sutterfield, A., Schein, J.E., Jones, S.J., Skalska, U., Smal, D.B.,
Schnur, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1122)
Straussberg, R.
Direct Submission
Submitted (09-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Sep 16, 2003 this sequence version replaced gi:13623362.
Contact: cgapbs-re@mail.nih.gov
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nigr.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Madsen, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Shantrop, S., Thomas, P.J., Touchman, J.W.,
Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 15 Row: d Column: 18
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 14150154.

FEATURES

Location/Qualifiers
 1..1122
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:10540 IMAGE:4025884"
 /tissue_type="Skin, melanotic melanoma."
 /clone_lib="NIH MGC 20"
 /lab_host="DH10B-R"
 /note="Vector: pOT87"
 1..1122
 /gene="MGC10540"
 /db_xref="LocusID:84313"
 41..571
 /gene="MGC10540"
 /codon_start=1
 /product="hypothetical protein MGC10540"
 /protein_id="AAH06282.1"
 /db_xref="GI:13623363"
 /db_xref="LocusID:84313"
 /translation="MAMSEFWQVRPPFFTLQPNVDTRQKLAACWCLVLSFRLH
 KQSMVMEASQSPFNKKLPVSIQIVLELRKKGNLWLDKSKSFLIMWR
 RPEWKLTYQVRSRQNNVSFTLYELTNGBDTEDEPHGLDEATLLRALQALQEH
 KAEIITVSDRGVKFP"

ORIGIN

Query Match 41.9%; Score 1092; DB 9; Length 1122;
 Best Local Similarity 100.0%; Pred. No. 8.5e-239;
 Matches 1092; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 CTGGGCTACTAGATGCGGATGAGTTTCAGTGGCGGTGCGAGTATCGCTTCCACCCCTT 168
 DB 28 CTGGGCTACTAGATGCGGATGAGTTTCAGTGGCGGTGCGAGTATCGCTTCCACCCCTT 87

QY 169 CTTTACGTTACACCGAATGTGACACTCGGCAGAGAGCTGGCGCTGTGCTCGCT 228
 DB 88 CTTTACGTTACACCGAATGTGACACTCGGCAGAGAGCTGGCGCTGTGCTCGCT 147

QY 229 GGTCTGTCTTCTGCGGCTGCAAAACAGTCCAGCATGAGTGTATGGAAGCTCAGGA 288
 DB 148 GGTCTGTCTTCTGCGGCTGCAAAACAGTCCAGCATGAGTGTATGGAAGCTCAGGA 207

QY 289 GAGCCGCTCTTCAACAGCTCAAGCTACAGCGAAGCTTCTGTGGAGTCCGATCCAGAT 348
 DB 208 GAGCCGCTCTTCAACAGCTCAAGCTACAGCGAAGCTTCTGTGGAGTCCGATCCAGAT 267

QY 349 TGTATTAGAGGAACCTGAGGAAGAAAGGAACTCGAGTGGTGGATAAGAGCAAGTCCAG 408
 DB 268 TGTATTAGAGGAACCTGAGGAAGAAAGGAACTCGAGTGGTGGATAAGAGCAAGTCCAG 327

QY 409 CTTCTGTATCATGTGGCGAGGCGAGAGAAATGGGGGAACTCATCTATCATGTGGGTTTC 468
 DB 328 CTTCTGTATCATGTGGCGAGGCGAGAGAAATGGGGGAACTCATCTATCATGTGGGTTTC 387

QY 469 CAGGATGCGCAGAACATCCGCTCTTACCTGTATGACTGACTAATGGGGAGACAC 528
 DB 388 CAGGATGCGCAGAACATCCGCTCTTACCTGTATGACTGACTAATGGGGAGACAC 447

QY 529 AGAGGATGAGGAGTTCCACGGGCTGATGAAGCCACTCTACTCGGGCTCTGACGGCCCT 588
 DB 448 AGAGGATGAGGAGTTCCACGGGCTGATGAAGCCACTCTACTCGGGCTCTGACGGCCCT 507

QY 589 ACAGCAGGAGCAACAGGCCAGATCATCTGTACGCGATGCGCGAGCGTCAAGTTCTT 648
 DB 508 ACAGCAGGAGCAACAGGCCAGATCATCTGTACGCGATGCGCGAGCGTCAAGTTCTT 567

QY 649 CTAGCAGGAGCCTGTCTCCCTTTACTTCTTACTCCCACTTTTCCAGGGCTTCAAAGG 708
 DB 568 CTAGCAGGAGCCTGTCTCCCTTTACTTCTTACTCCCACTTTTCCAGGGCTTCAAAGG 627

QY 709 AGACAGACCCAGTGTGCTCCCAAAAGACTGTGACTTCCACAGACTCAAAAGGACTC 768
 DB 628 AGACAGACCCAGTGTGCTCCCAAAAGACTGTGACTTCCACAGACTCAAAAGGACTC 687

QY 769 CAGTCTGAGGCTGGGACCTGGGATGGTTTCTCACACCCCATATGTCTGTCCCTTGG 828
 DB 688 CAGTCTGAGGCTGGGACCTGGGATGGTTTCTCACACCCCATATGTCTGTCCCTTGG 747

QY 829 ATAGGCTGAGGCTGAAGACACAGGAGAGAAATATGTGCTTCTTCTCGCCCTACCTCCTTT 888
 DB 748 ATAGGCTGAGGCTGAAGACACAGGAGAGAAATATGTGCTTCTTCTCGCCCTACCTCCTTT 807

QY 889 CCATCTTAGACTGTCTTGTGAGCCAGGGTCTGTAAACCTGACACTTATATGTGTTTACA 948
 DB 808 CCATCTTAGACTGTCTTGTGAGCCAGGGTCTGTAAACCTGACACTTATATGTGTTTACA 867

QY 949 CATGTAAGTACATACACATGCGCTGCGAGCAGATGCTTCTGTCTCTCTCTCTCTCCAC 1008
 DB 868 CATGTAAGTACATACACATGCGCTGCGAGCAGATGCTTCTGTCTCTCTCTCTCTCCAC 927

QY 1009 CCCTTTAGCTGTGTTGCTTCTCAGGCTGGTGTGATCCTTCTTAGGGATGGG 1068
 DB 928 CCCTTTAGCTGTGTTGCTTCTCAGGCTGGTGTGATCCTTCTTAGGGATGGG 987

QY 1069 GGAAGCCCTGGCTGACGAGCAGCTTCCAGGCAATATGAAGATAGGAGGCCACGGGCTG 1128
 DB 988 GGAAGCCCTGGCTGACGAGCAGCTTCCAGGCAATATGAAGATAGGAGGCCACGGGCTG 1047

QY 1129 GCAGTGAGAGGTGTGGCCCCACACCGATTATGATATTAATAATCTCAACTCCCAAAAAA 1188
 DB 1048 GCAGTGAGAGGTGTGGCCCCACACCGATTATGATATTAATAATCTCAACTCCCAAAAAA 1107

QY 1189 AAAAAAATAA 1200
 DB 1108 AAAAAAATAA 1119

RESULT 5
 AR339439
 LOCUS 1243 bp DNA linear PAT 17-AUG-2003
 DEFINITION Sequence 930 from patent US 6569662.
 ACCESSION AR339439
 VERSION AR339439.1 GI:33726296
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 UNCLASSIFIED.
 REFERENCE 1 (bases 1 to 1243)
 AUTHORS Tang, Y.T., Zhou, P. and Drmanac, R.T.
 TITLE Nucleic acids and polypeptides
 JOURNAL Patent: US 6569662-A 930 27-MAY-2003;
 FEATURES Location/Qualifiers
 source
 1..1243
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN
 Query Match 41.3%; Score 1076.2; DB 6; Length 1243;
 Best Local Similarity 99.7%; Pred. No. 3.5e-235;
 Matches 1078; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 113 GCTCTACGATGGCGATGAGTTTCAGTGGCGGTGGCAGTATCGCTTCCACCCCTCTTT 172
 DB 163 GCGTCGACGATGGCGATGAGTTTCAGTGGCGGTGGCAGTATCGCTTCCACCCCTCTTT 222

QY 173 AGCTTACACCGAATGTGACACTCGGAGAGAGAGCTGGCCCTGTGCTGTGCTGTGTC 232
 DB 223 AGCTTACACCGAATGTGACACTCGGAGAGAGAGCTGGCCCTGTGCTGTGCTGTGTC 282

QY 233 CTGTCTCTTCTGCGGCTTGCACAAACAGTCCAGATGAGTGTGATGGAAGCTCAGGAGAGC 292
 DB 283 CTGTCTCTTCTGCGGCTTGCACAAACAGTCCAGATGAGTGTGATGGAAGCTCAGGAGAGC 342

QY	293	CGCTCTTTCAACAACTCAAGCTACAGCGAAAGCTTCTGTGGAGTCGATCCAGATTGTA	352
Db	343	CGCTCTTTCAACAACTCAAGCTACAGCGAAAGCTTCTGTGGAGTCGATCCAGATTGTA	402
QY	353	TTAGAGGAACCTGAGGAAGAAAGGAACCTCGAGTGGTTGGATAAGACGAAGTCCAGCTTC	412
Db	403	TTAGAGGAACCTGAGGAAGAAAGGAACCTCGAGTGGTTGGATAAGACGAAGTCCAGCTTC	462
QY	413	CTGATCATGTGGCGAGGCGCAGAGAAGTGGGGAAACTCATCTATCAGTGGTTCACGAG	472
Db	463	CTGATCATGTGGCGAGGCGCAGAGAAGTGGGGAAACTCATCTATCAGTGGTTCACGAG	522
QY	473	AGTGCCACAGAACTCGCTCTTTTACCTGTATGAACCTGACTAAATGGGGAAGACACAG	532
Db	523	AGTGCCACAGAACTCGCTCTTTTACCTGTATGAACCTGACTAAATGGGGAAGACACAG	582
QY	533	GATGAGGAGTTCACCGGCTGGATGAAGCCACTCTACTGCGGGCTCTGACGSCCTACAG	592
Db	583	GATGAGGAGTTCACCGGCTGGATGAAGCCACTCTACTGCGGGCTCTGACGSCCTACAG	642
QY	593	CAGGAGCAGAGCCGAGATCATCTGTACCGATGGCCGAGGCGTCAAGTCTTCTAG	652
Db	643	CAGGAGCAGAGCCGAGATCATCTGTACCGATGGCCGAGGCGTCAAGTCTTCTAG	702
QY	653	CAGGACCTGTCTCCCTTTACTTCTTACTCTCCCTCCACTTTCCAGGGCTTTCAAAGGAGAC	712
Db	703	CAGGACCTGTCTCCCTTTACTTCTTACTCTCCCTCCACTTTCCAGGGCTTTCAAAGGAGAC	762
QY	713	AGACCCAGTGTCCCCCAAGACTGGATCTGTGACTCCACGACTCAAAAGGACTCCAGT	772
Db	763	AGACCCAGTGTCCCCCAAGACTGGATCTGTGACTCCACGACTCAAAAGGACTCCAGT	822
QY	773	CCTGAAGCGTGGGACTGGGATGGGTTCTCACAACCCGATATGTCTGTCCTTGGATAG	832
Db	823	CCTGAAGCGTGGGACTGGGATGGGTTCTCACAACCCGATATGTCTGTCCTTGGATAG	882
QY	833	GCTGAGGCTGAAGCACACGAGGAGAAATATGTGCTTCTCGCCCTACCTCTTTCCCA	892
Db	883	GCTGAGGCTGAAGCACACGAGGAGAAATATGTGCTTCTCGCCCTACCTCTTTCCCA	942
QY	893	TCCTAGACTGCTTTGAGCAGGGGCTGTGAACCTGACACTTTATATGTTGTTCAACATG	952
Db	943	TCCTAGACTGCTTTGAGCAGGGGCTGTGAACCTGACACTTTATATGTTGTTCAACATG	1002
QY	953	TAAGTACATACACATGCGGCTGACGACATGCTTCTGCTCTCTCTCTCCACCCCT	1012
Db	1003	TAAGTACATACACATGCGGCTGACGACATGCTTCTGCTCTCTCTCTCTCCACCCCT	1062
QY	1013	TTAGCTGCTGTTGCTCTCTCTCAGGCTGGTGGATCTCTTCTAGGGATGGGGAA	1072
Db	1063	TTAGCTGCTGTTGCTCTCTCTCAGGCTGGTGGATCTCTTCTAGGGATGGGGAA	1122
QY	1073	GCCCTGGCTGACGAGGCTTCCAGGCAATATGAAGATAGGAGGCCACCGGCTGGCAG	1132
Db	1123	GCCCTGGCTGACGAGGCTTCCAGGCAATATGAAGATAGGAGGCCACCGGCTGGCAG	1182
QY	1133	TCGAGAGGTGGCCCCCACACCGATTATGATATTAATCTCAACTCCCAAAAAA	1192
Db	1183	TCGAGAGGTGGCCCCCACACCGATTATGATATTAATCTCAACTCCCAAAAAA	1242
QY	1193	A	1193
Db	1243	A	1243
RESULT 6			
E50594			
LOCUS	E50594	1087 bp	DNA linear PAT 27-AUG-2002
DEFINITION	Human protein and cDNA [3].		
ACCESSION	E50594		
VERSION	E50594.1 GI:22555017		
KEYWORDS	JP 2001161368-A/14.		

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 1087)
AUTHORS	Kato,S. and Saeki,M.
TITLE	Human protein and cDNA [3]
JOURNAL	Patent: JP 2001161368-A 14 19-JUN-2001; SCIENCE & TECH AGENCY
COMMENT	OS Homo sapiens (human) PN JP 2001161368-A/14 PD 19-JUN-2001 PF 06-DEC-1999 JP 1999346864 PI SEIJI KATO,MIHORO SAEKI PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/10//
PC C12P21/02,C12P21/08,C12N15/00,C12N5/00	
CC	
FH	Key Location/Qualifiers
FT	CDS (28)..(558).
FEATURES	
source	1..1087
/organism="Homo sapiens"	
/mol_type="genomic DNA"	
/db_xref="taxon:9606"	
ORIGIN	
Query Match 41.2%; Score 1073; DB 6; Length 1087;	
Best Local Similarity 100.0%; Pred. No. 1.9e-234;	
Matches 1073; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	109 CTGGGCTACTACGATGGCGATGAGTTTCGAGTGGCGGTCGAGTATCGCTTCCACCCCTT 168
Db	15 CTGGGCTACTACGATGGCGATGAGTTTCGAGTGGCGGTCGAGTATCGCTTCCACCCCTT 74
QY	169 CTTTACGTTTACAACCGAATGTGGACACTCGGCAGAACGAGCTGGCGGCTGGTGTCTCGCT 228
Db	75 CTTTACGTTTACAACCGAATGTGGACACTCGGCAGAACGAGCTGGCGGCTGGTGTCTCGCT 134
QY	229 GGTCTGTCTTCTGCGCCCTGCGACAAACAGTCCAGCATGACGGTGATGGAAGCTCAGGA 288
Db	135 GGTCTGTCTTCTGCGCCCTGCGACAAACAGTCCAGCATGACGGTGATGGAAGCTCAGGA 194
QY	289 GAGCCCGCTCTTCAACACGCTCAAGCTACAGGAAAGCTTCTGTGGAGTGGATCCAGAT 348
Db	195 GAGCCCGCTCTTCAACACGCTCAAGCTACAGGAAAGCTTCTGTGGAGTGGATCCAGAT 254
QY	349 TGTATTAGAGGAACCTGAGGAAGAAAGGAACTCTCGAGTGGTTGGATAAGAGCAAGTCCAG 408
Db	255 TGTATTAGAGGAACCTGAGGAAGAAAGGAACTCTCGAGTGGTTGGATAAGAGCAAGTCCAG 314
QY	409 CTTCTGTATCATGTGGCGGAGGCGCAGAAAGATGGGGGAAACTCATCTATCAGTGGGTTTC 468
Db	315 CTTCTGTATCATGTGGCGGAGGCCAGAAAGATGGGGGAAACTCATCTATCAGTGGGTTTC 374
QY	469 CAGGAGTGGCCAGAACAACTCCGCTTTTACCTCTTATGAACTGACTAATGGGGAAGACAC 528
Db	375 CAGGAGTGGCCAGAACAACTCCGCTTTTACCTCTTATGAACTGACTAATGGGGAAGACAC 434
QY	529 AGAGGATGAGGAGTTTCCACGGGCTGGATGAAGCCACTCTACTCGGGGCTCTGACGGCCCT 588
Db	435 AGAGGATGAGGAGTTTCCACGGGCTGGATGAAGCCACTCTACTCGGGGCTCTGACGGCCCT 494
QY	589 ACAGCAGAGCACAAGGCCGAGATCATCTGTCTAGCGATGGCCGAGGCGTCAAGTTCTT 648
Db	495 ACAGCAGAGCACAAGGCCGAGATCATCTGTCTAGCGATGGCCGAGGCGTCAAGTTCTT 554
QY	649 CTAGCAGGACCTGTCTCCCTTTTACTTCTTACCTCCACCTTTTCCAGGGCTTTCAAAAGG 708
Db	555 CTAGCAGGACCTGTCTCCCTTTTACTTCTTACCTCCACCTTTTCCAGGGCTTTCAAAAGG 614
QY	709 AGACAGACCCAGTGTCTCCCAAGAGCTGGATCTGTGACTCTCCACGAGACTCAAAAGGACTC 768

Db 1035 GCAGTGAGAGGTGTGGCCCCCACACCGATTATGATATTAATAATCTCAACTCCC 1087

RESULT 8

BC029181

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

BC029181 1112 bp mRNA linear ROD 29-JUN-2004
Mus musculus DNA segment, Chr 11, Wayne State University 68,
expressed, mRNA (cDNA clone MGC:35627 IMAGE:4240500), complete cds.
BC029181
MGC.
BC029181.1 GI:20809732
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1112)
Strausberg,R.D., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,P.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Faney,J., Helton,E., Kettner,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzyzanski,M.I., Skaleka,U., Smalish,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1112)
Strausberg,R.
Direct Submission
Submitted (01-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louleeged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 61 Row: m Column: 3
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 25092661.
Location/Qualifiers
1. .1112
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:35627 IMAGE:4240500"
/tissue_type="Kidney, normal, 5 month old male mouse."

FEATURES

source

gene

CDS

/clone.lib="NCI CGAP_Kid14"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .1112
/gene="DilWsu68e"
/db_xref="LocusID:28084"
/db_xref="MGI:106354"
13. -543
/gene="DilWsu68e"
/codon_start=1
/product="DNA segment, Chr 11, Wayne State University 68,
expressed"
/protein_id="AAH29181.1"
/db_xref="GI:20809733"
/db_xref="LocusID:28084"
/db_xref="MGI:106354"
/translation="MAMSFEPWQYRFPFTLPQNVDTROKQLAAMCSLVLSFRLH
KQSMVTMEAESPLFNKVLKIPVSIQIVLEELKKGNLEWLDKNSFLIMWR
RPEWKGILYQWVSRSQNNVSFTLYELTSGEDTEDEFHGLDEATLLRALQALQOEH
KAEIITVSDGRGVKFF"

ORIGIN

Query Match 30.2%; Score 786.8; DB 10; Length 1112;
Best Local Similarity 87.9%; Pred. No. 5.7e-169;
Matches 963; Conservative 0; Mismatches 117; Indels 16; Gaps 9;
QY 110 TGGGCTACTAGATGGCGATGAGTTTCGAGTGGCGTGCAGTATCGCTTCCACCCCTTC 169
Db 1 TGGGCTACTAGATGGCGATGAGTTTCGAGTGGCGTGCAGTATCGCTTCCGCGCTTC 60
QY 170 TTTAGCTTACACCAATGTGGACACTCGGCAGAACAGCTGGCGCCCTGGTGTCTCGCTG 229
Db 61 TTTAGCTTACACCGAAGTGGACACTCGGCAGAACAGCTGGCGCCCTGGTGTCTCTGTG 120
QY 230 GTCTGTCTTCTGCGGCTGCAACACAGTCCAGCTGACGCTGATGAGTGAAGTCCAGAG 289
Db 121 GTTCTGTCTTCTGCGGCTGCAACACAGTCCAGCTGACGCTGATGAGTGAAGTCCAGAG 180
QY 290 AGCCGCTCTTCAACACGTCAGCTACAGCGAAGCTTCTGTGGAGTCCAGTCCAGATT 349
Db 181 AGCCGCTCTTCAACACGTCAGCTACAGCGAAGCTTCTGTGGAGTCAATTGAGATT 240
QY 350 GTATTAGAGGAACCTGAGGAAGAAAGGAAACCTCGAGTGGTGGATAAGAGCAAGTCCAGC 409
Db 241 GTATTAGAGGAACCTGAGGAAGAAAGGAAACCTCGAGTGGTGGATAAGAAATAGTCTAGC 300
QY 410 TTCCTGATCATGTGGCGAGGCGCAGAACAAATGGGGAAACTCATCTATCAGTGGGTTC 469
Db 301 TTCCTGATCATGTGGCGAGGCGCAGAACAAATGGGGAAACTCATTTACCAATGGGTTC 360
QY 470 AGGAGTGCCAGAACCAACTCCGTCTTTTACCTGTATGAACCTGACTTAATGGGGAACACACA 529
Db 361 AGGAGTGCCAGAACCAACTCTGTGTCTGTCTGTATGAACTGACGAGTGGGGAACACACA 420
QY 530 GAGGATGAGGAGTTCACGGGCTGGATGAAGCAACTCTACTGCGGGCTCTGAGGCCCTTA 589
Db 421 GAGGATGAAGAGTTCACGGGCTGGATGAAGCAACTCTACTGCGGGCTCTGAGGCCCTTA 480
QY 590 CAGCAGGAGCACAAGGCCGAGATCACTGTCAGCGATGGCCGAGGCGTCAAGTCTTC 649
Db 481 CAGCAGGAGCACAAGGCCGAGATCACTGTCAGCGATGGCCGAGGCGTCAAGTCTTC 540
QY 650 TAGCAGGACCTGTCTCCCTTTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 709
Db 541 TAGCAGGACCTGTCTCCCTTTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 599
QY 710 GACAGACCCAGTGTCCCCCAAGAGCTGGATCTGTGACTTCCACAGACTCAAAAGGACTCC 769
Db 600 GACAGACCTAGTGT-CCCCACAGACTGGATCTGTCTCACTCTCTCCAGACTC-AAAGGGCTCA 657
QY 770 AGTCTTGAAGCTGGGAGCTGGGATGGGTTTCTCA-CACCCCATATCTGTCTGCTCTGG 828
Db 658 AGTCTTGAAGCTCAGACTTAGGGATGTTTCCACATAACCCCTTTATTCTGTCTCTCTGG 717

```

Qy 829 ATAGGTGAGGCTGAAGCACCCAGGGAGAAAAATATGTCTTCTTCTGCGCCCTACCTCCTTT 888
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
718 ATAGAGTAAGCCACACAGTCCACCGGAGAAAGGATGGTTCTT---GCCCTACCTCCTTT 774
Qy 889 CCATCTAGACTGCTTTCAGCCAGGCTGTAAACCTGACACTTATATGTGTTCACA 948
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
775 CCATCTTACTGTGCTCCCGAGCCAGGGCTATAAACCTGACACTTATATGTGTTCACA 834
Qy 949 CATGTAAGTACATACACATGCGCTGCGAGCACATGCTTCTGTCTCCCTCCTCCTCCAC 1008
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
835 CACATAAGTACACACCACT-ACATGCGCTGCGACAGCTTCTTCTCCTCCTCCAC 892
Qy 1009 CCTTTAGCTGTGTTCCTCCTTCTCAGGCTGTGTGTGATCTTCTTCTTCTTCTTCTTCT 1068
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
893 TCCTTTAGCTGTGTTCCTCCTTCTCAGGCTGTGTGTGATCTTCTTCTTCTTCTTCTTCT 951
Qy 1069 GGAAGCCCTGCTGAGGAGCTTCCAGGCAATATGAGATAGGAGGCC---ACGGG 1124
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
952 GGAAGCCCTGCTGAGGAGCTTCCAGGCGCGGCGAGTATGAGATAGGAGGCCCGAGAGGG 1011
Qy 1125 CTGGCAGTGAGGCTGCGCCCAACCGATTTATGATTTAAATCTCAACTCCCAA 1184
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1012 CTGTGTGTGAGAG--GCGGGCCACTCTGATTTATGATTTAAATCTTCTACTCCAA 1069
Qy 1185 AAAAAAAAAAAAAA 1200
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1070 AAAAAAAAAAAAAA 1085

```

RESULT 9

```

AL157902
LOCUS      Human DNA sequence from clone RP4-675C20 on chromosome lp13.2.
DEFINITION Contains the 3' end of the MAN1A2 gene for mannosidase alpha 1A2, a
            pseudogene similar to predicted fly, worm and yeast genes, ESTs,
            STSs and GSSs, complete sequence.
ACCESSION  AL157902.6 GI:9581579
VERSION    HTG; MAN1A2; mannosidase.
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 116397)
AUTHORS    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLES     Coville, G.
JOURNAL    Direct Submission
            Submitted (24-AUG-2000) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
            requests: clonerequest@sanger.ac.uk
            On Jul 28, 2000 this sequence version replaced gi:9437233.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence has been finished according to sequence map criteria
            as follows. An attempt is made to resolve all sequencing problems,
            such as compressions and repeats, but not necessarily within known
            annotated human repeat sequence elements (e.g. Alu). Where the
            sequence is ambiguous, there is an annotation using the 'unsure'
            feature key.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group. Further information can be found at
            http://www.sanger.ac.uk/HGP/Chri
            RP4-675C20 is from the library RPCI-4 constructed at the Roswell

```

Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/>
 VECTOR: pCVPAC2
 This sequence is the entire insert of clone RP4-675C20 The true left end of clone RP11-42121 is at 81459 in this sequence. The true right end of clone RP11-188D8 is at 37690 in this sequence.

FEATURES

source

```

1..116397
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="p13.2"
/clone="RP4-675C20"
/clone_lib="RPCI-4"
72..786
/note="HAL1 repeat: matches 170..910 of consensus"
repeat_region
1756..2037
/note="AluX repeat: matches 1..298 of consensus"
repeat_region
2074..2311
/note="L1MD2 repeat: matches 5960..6192 of consensus"
repeat_region
2456..2486
/note="Alu repeat: matches 56..87 of consensus"
repeat_region
2873..3064
/note="L1ME repeat: matches 5320..5513 of consensus"
repeat_region
3696..6108
/note="TIGER1 repeat: matches 20..2418 of consensus"
6689..42414
/gene="MAN1A2"
join(<6689..6804,10305..10524,12924..13096,16397..16512,
36367..42414)
/gene="MAN1A2"
/product="dJ675C20.1 (mannosidase, alpha, class 1A, member 2)"
/note="match: cDNAs: Em:U03458 Em:U03457 Em:X74837
Em:U04301 Em:Y12503 Em:U04299 Em:AF027156 Em:AK001970
match: ESTs: Em:AI122755 Em:AI084575 Em:N41869 Em:AI379549
Em:AI197606 Em:AI861966 Em:AA544899 Em:AI358786 Em:B62885
Em:R22905 Em:AI961563 Em:R31736 Em:AA236264 Em:AA856747
Em:AI506935 Em:AI119741 Em:AI566231 Em:AW021115
Em:AA478661 Em:AI169902 Em:AI397108 Em:AI169903
Em:AA748618 Em:AI355944 Em:AI149583 Em:AI378391
Em:AA578856 Em:AI056108 Em:R23007 Em:AA013940 Em:AA410832
Em:R71509 Em:R80653 Em:AA235544 Em:AA550229 Em:AA832229
Em:Y10746 Em:AI810315 Em:AI170182 Em:AI472133 Em:AI170436
Em:AI004952 Em:AA874194 Em:R77137 Em:245667 Em:AI133955
Em:AV325150 Em:AA830923 Em:R77136 Em:AI804479 Em:AW333036
Em:R22121 Em:D31578 Em:DB1772 Em:R78501 Em:R8060
Em:R85259 Em:AW558740 Em:AA323243 Em:W33047 Em:R82762
Em:AA993022 Em:R64469 Em:R68908 Em:AI55323 Em:N30111
Em:H01427 Em:AI567623 Em:H01428 Em:R31041 Em:AA988792
Em:AW164558 Em:N92139 Em:AA507630 Em:R20993 Em:N27386
Em:NA0798 Em:R78550 Em:AI151416 Em:AA909450 Em:AA507636
Em:AI817604 Em:AA406590 Em:AI025868 Em:AI717735
Em:AW204118 Em:AI457883 Em:AI093841 Em:R20443 Em:AW074094
Em:D61957 Em:R22064 Em:AI346166 Em:AW235689 Em:AA394233
Em:R24488 Em:AI192314 Em:AI633433 Em:R36410 Em:AA448826
Em:W37584 Em:AA478626 Em:R65698 Em:R65699 Em:AA478864
Em:AI337574 Em:AW573122 Em:AI823859 Em:AW338072
Em:AA224016 Em:R80450 Em:H97940 Em:R12222 Em:AI16939
Em:AA615144 Em:AA223923 Em:AV37431 Em:AV335409
Em:AI289196 Em:AA304684 Em:AA628731 Em:AA399331 Em:R68059
Em:AA307758 Em:AI289102 Em:W19691 Em:AI888230 Em:AI417585
Em:N32009 Em:N30624 Em:AI889444 Em:N40905 Em:D18568
Em:R23372"
/evidence=not_experimental
join(<6689..6804,10305..10524,12924..13096,16397..16512,
36367..42499)
/gene="MAN1A2"
/note="match: proteins: Tr:O02773 Sw:P53624 Sw:P53625
Tr:O18498 Sw:P39098 Sw:P33908 Tr:Q9UJ13 Sw:P45700
Tr:O60476 Sw:P45701"
/codon_start=3

```

CDS

```

/evidence=not_experimental
join(<6689..6804,10305..10524,12924..13096,16397..16512,
36367..42499)
/gene="MAN1A2"
/note="match: proteins: Tr:O02773 Sw:P53624 Sw:P53625
Tr:O18498 Sw:P39098 Sw:P33908 Tr:Q9UJ13 Sw:P45700
Tr:O60476 Sw:P45701"
/codon_start=3

```

/evidence=not experimental
/product="d3675C20.1 (mannosidase, alpha, class 1A, member 2)"
/protein_id="CAC10451.1"
/db_xref="GI:10443041"
/db_xref="GOA:O60476"
/db_xref="UniProt/Swiss-Prot:O60476"
/translation="HTSVGLGDSFYVLLKAWLMSDKTDHEARKMYDDAIEATEKHL
IKSGRGTLFTGEWKNHLEKMGHLACFAGMGALGADGSRADKAGHYLGLGAEIAR
TKHSYDRTALKLPESPKFDGPAVAVRQAEKYIILRPVETWYWLWRFTHDPYR
RQWGEALALEIKYCRVNGSGVSDYTSSTPHDDVQQSFLEATLKLKLLFSGDD
LLPLDHWVFNEAHLPLVHLANTLTLGNPAVR"
7043. .7347
repeat_region
/note="AluSx repeat: matches 3. .307 of consensus"
7506. .7733
repeat_region
/note="MIR repeat: matches 1. .262 of consensus"
8136. .8179
repeat_region
/note="L2 repeat: matches 2657. .2700 of consensus"
8363. .8480
repeat_region
/note="L2 repeat: matches 2584. .2710 of consensus"
complement(10218. .10720)
misc_feature
/note="match: GSS: Em:AQ664370"
11953. .12065
repeat_region
/note="L2 repeat: matches 2592. .2695 of consensus"
complement(13408. .13965)
misc_feature
/note="match: GSS: Em:AQ310773"
13564. .14014
misc_feature
/gene="MAN1A2"
/note="match: GSS: Em:AQ876874"
15669. .15813
repeat_region
/note="MIR repeat: matches 93. .259 of consensus"
16746. .16903
repeat_region
/note="MIR repeat: matches 13. .175 of consensus"
17228. .17336
repeat_region
/note="L2 repeat: matches 2627. .2749 of consensus"
17869. .18611
repeat_region
/note="L1M4 repeat: matches 7182. .7960 of consensus"
19003. .19046
repeat_region
/note="11 copies 4 mer at 81% conserved"
19168. .19221
repeat_region
/note="2 copies 27 mer 100% conserved"
20219. .20318
repeat_region
/note="L2 repeat: matches 2418. .2521 of consensus"
20642. .20926
repeat_region
/note="HAL1 repeat: matches 467. .763 of consensus"
24238. .24350
repeat_region
/note="L2 repeat: matches 2629. .2746 of consensus"
24497. .24864
repeat_region
/note="THE1C repeat: matches 1. .371 of consensus"
24894. .25015
repeat_region
/note="FLAM_C repeat: matches 2. .123 of consensus"
25042. .25148
repeat_region
/note="L2 repeat: matches 2621. .2742 of consensus"
25214. .25516
repeat_region
/note="L2 repeat: matches 2440. .2748 of consensus"
25877. .26061
repeat_region
/note="MERSA repeat: matches 1. .188 of consensus"
26291. .26338
repeat_region
/note="24 copies 2 mer ac 87% conserved"
26293. .26340
repeat_region
/note="12 copies 4 mer acac 87% conserved"
26549. .26738
repeat_region
/note="L1M82 repeat: matches 5903. .6120 of consensus"
26799. .26887
repeat_region
/note="L1M83 repeat: matches 6087. .6175 of consensus"
26892. .27031
repeat_region
/note="L1M82 repeat: matches 5749. .5896 of consensus"
27623. .27930
repeat_region
/note="AluSx repeat: matches 1. .309 of consensus"
28659. .28943
repeat_region
/note="MER2 repeat: matches 1. .345 of consensus"
29210. .29489
repeat_region
/note="L1M82 repeat: matches 5841. .6166 of consensus"

misc_feature
complement(29909. .30339)
/note="match: GSS: Em:AQ379387"
30087. .31494
repeat_region
/note="L1M8c repeat: matches 274. .1745 of consensus"
32097. .34934
repeat_region
/note="L1PB1 repeat: matches 2354. .5224 of consensus"
34905. .35728
repeat_region
/note="L1PB1 repeat: matches 5336. .6155 of consensus"
35756. .36102
repeat_region
/note="L1M4 repeat: matches 3703. .4100 of consensus"
36401. .36820
misc_feature
/gene="MAN1A2"
/note="match: GSS: Em:AQ093391"
Query Match 26.8%; Score 697.6; DB 9; Length 116397;
Best Local Similarity 80.3%; Pred. No. 3.3e-148;
Matches 925; Conservative 0; Mismatches 199; Indels 28; Gaps 8;
QY 96 AATTCCGACGAGCTGGGCTACTACGATGGCGATGAGTTTCGAGTGGCGGTGGCAGTATC 155
Db AAAGTGGGATTTCTCTGGGCTATTATGATGGTGACGAATTTCAAGTGGTTGTGGCAGTATC 62990
QY 156 GCTTCCCAACCTTTCTTTACGTTTAAACCGAATGTGGACACTCGGCGAGAACGAGCTGGCGG 215
Db GCTTTCCGCGC--CTTTACATTACAGCTGAACGTGGCCACTTTGGCAGAAGCAGCTGGCCA 63048
QY 216 CTTGTGTCTCGTGTCTGTCTCTTCGCGCTGCACAAACAGTCCAGCATGACGCTGA 275
Db CTTGTGTGTGTGTGTCTGTCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 63048
QY 276 TGGAGCTCCAGGAGGCGGCTCTTCAACAACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 335
Db TGGATGCTCAGGAGATCTGCTCTTCAGCAACATCAAGCT--GTGGAAGCTTCCTGTGG 63165
QY 336 AGTCGATCCAGATTGTATTAGAGAACTAGAGAAAGGAAAGCACTCGAGTGGTGGATA 395
Db GATCAATCCAGTTGTATTAGAGAACTAGAGAAAGTGGAACTTACAGTGGCTGGATA 63225
QY 396 AGAGCAAGTCCAGCTTCCTGATCATGTGCGGAGCCAGAGAAATGGGGAAACTCATCT 455
Db AGAGCAAGTCTAGTTTCTCTAATCATGTGGGAGGCCAGAGAAATGGGGAAACTCATCT 63285
QY 456 ATCAGTGGGTTTCCAGGAGTGGCCAGAAACAATCCGTCCTTTTACCCCTGTATGAACCTACT 515
Db ATCAGTGGGTTTCCAGGAGTGGCCAGAAACAATCCGTCCTTTTACCCCTGTATGAACCTACT 63345
QY 516 ATGGGAAAGACACAGAGATGAGGAGTTCCAGGCTGGATGAAGCCACTTACTTGGCGG 575
Db ATGGGAAAGACATAGAGAAATGAGGTTTCCACGGACTAAAGGAGGC----CTTCTGTGGG 63401
QY 576 CTCTGAGGCGCTTCCAGCAGGAGCACAAAGGCGGAGATCATCACTGT-----C 622
Db CTCTGAGGCGCTTCCAGCAGGAGCACAAAGGCGGAGATCATCACTGT-----C 63461
QY 623 AGCGATGGCCGAGGCGTCAAGTTCTTTAGCAGGAGACCTGTCTCCCTTTTACTTTTACT 682
Db AGTATGCTGAGGTGTCAAGTTCATCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 63521
QY 683 CCCACCTTTCCAGGCGCTTTCAAAAGGAGACAGACCCAGTGTCCCAAGACTGATCTG 742
Db CCCACCTTTCCAGGCGCTTTTCAGGAGGAGACAGACTCGGCTG--GACCACACATTTGAACCTG 63580
QY 743 TGACTCCACACAGACTCAAAAGGAGCTCCAGTCTGAAGCTGGGAGCTGGGATGGGTTTC 802
Db TGAATCCACACAGACTCAAAAGGAGCTCCAGTCTGAAGCTGGGAGCTGGGATGGGTTTC 63640
QY 803 TCACACCCCATATGTCTGTCTCCCTTGGATAGGCTGAGGCTGAAGCACACGAGGAGAAAAAT 862
Db TCATACACAAATGTCTGTCTCCCTTGGACAGAGGTGAGGCTGAGGCACTCCAGAGAAAAATG 63700
QY 863 GTGCTCTTCTTC--GCCCTACCTCTCTTCCCTCCTAGACTGTCTTGGACCCAGGCTGTGT 921
Db CTTTCAAGGCTTGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTGT 63760

QY 922 AACCTGACACTTATATGTTTACACATGTAAGTATACATACACATGCGCTGACGCA 981
 |||||
 Db 63761 AACCTAACACTTATGTTGTTTACATATGTAAGTTTACACATGCACTC--GCA 63817
 |||||
 QY 982 CATGCTTCTGTCCTCTCCCTCCACCCCTTTAGCTGCTGTTGCTCCCTCTCAGGCT 1041
 |||||
 Db 63818 CATGCTTCTGTCCTCCCTCATCTTCTATCCCTTAGCTGCTGTTGCTCCCTCTCAGGCT 63877
 |||||
 QY 1042 GGTGCTGGATCTCTCTAGGGATGGGGAAGCCCTGGCTGACAGGAGGCTTCCAGGCAA 1101
 |||||
 Db 63878 GCTGCTGGATCTTCTCTAGGATGTGGGAG-CTTGGCTGACGAGGAGCATCCAGCAA 63936
 |||||
 QY 1102 TATGAAGATAGAGGCGCCAGGCGCTGGCAGTGAAGGTGTGGCCGCCACACCGATTATG 1161
 |||||
 Db 63937 TAGAAGATAGAGGCGCCATGGCGCTGGCACTGACAGGAGGCGCCCAACATTGATTATG 63996
 |||||
 QY 1162 ATATTAAATCTCACTCCCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1221
 |||||
 Db 63997 TTATTAAATCTCACTGCGCCCACTTCAGACAAATAAAAAAATAAAAAAATAAAAAA 64056
 |||||
 QY 1222 TCGAGAACTAGT 1233
 |||||
 Db 64057 AATAATTCAGT 64068
 |||||

RESULT 10
 AC100793
 LOCUS
 DEFINITION Homo sapiens chromosome 17, clone CTD-3193K9, complete sequence.
 AC100793
 VERSION
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 134465)
 AUTHORS Birren,B., Nussbaum,C. and Lander,E.
 TITLE Homo sapiens chromosome 17, clone CTD-3193K9
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 134465)
 Birren,B., Linton,L., Nussbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
 Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Heatford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
 Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
 Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,C., Nicol,R., Norbu,C.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
 Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 134465)
 AUTHORS Birren,B., Nussbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
 Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
 Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Heatford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
 Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
 Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,C., Nicol,R., Norbu,C.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
 Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 134465)

REFERENCES
 AUTHORS Birren,B., Nussbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
 Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
 Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,

TITLE
JOURNALREFERENCE
AUTHORS

4 (bases 1 to 134465)
 Birren,B., Nussbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
 Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
 Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
 Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
 Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
 Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
 Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
 Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
 Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
 Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
 Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
 Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
 Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL

COMMENT

Submitted (15-JAN-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jan 15, 2003 this sequence version replaced gi:24415443.
 All repeats were identified using RepeatMasker:
 Smit, A.P.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Center project name: L20710

Center clone name: 3193_K_9

FEATURES

source

Location/Qualifiers
 1..134465
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17"
 /clone="CTD-3193K9"
 /clone_lib="CITD2 Human BAC"

repeat_region

1..88
 /rpt_family="AluJo/FRAM"

repeat_region

94..389
 /rpt_family="AluSx"

repeat_region

complement(479..551)
 /rpt_family="Tigger5"

repeat_region

575..694
 /rpt_family="MER21B"

repeat_region

1296..1431
 /rpt_family="AluJb"

repeat_region

1432..1728
 /rpt_family="AluSx"

repeat_region

1731..2062

repeat_region /rpt_family="AluJb"
2400..2547
repeat_region /rpt_family="L2"
2613..2661
repeat_region /rpt_family="(TA)n"
complement(2670..2789)
repeat_region /rpt_family="FLAM_C"
complement(2795..3122)
repeat_region /rpt_family="AluY"
6180..6232
repeat_region /rpt_family="L2"
complement(6233..6530)
repeat_region /rpt_family="AluY"
6531..6585
repeat_region /rpt_family="L2"
complement(6586..6887)
repeat_region /rpt_family="AluSc"
6888..7159
repeat_region /rpt_family="L2"
7875..7912
repeat_region /rpt_family="GC_rich"
7957..8044
repeat_region /rpt_family="(CGGGG)n"
9590..9691
repeat_region /rpt_family="MER91A"
10084..10105
repeat_region /rpt_family="GC_rich"
10488..10589
repeat_region /rpt_family="MIR"
12437..12588
repeat_region /rpt_family="MIR3"
12868..12900
repeat_region /rpt_family="(CCCCCT)n"
13333..13370
repeat_region /rpt_family="GC_rich"
14420..14475
repeat_region /rpt_family="GA_rich"
14721..14744
repeat_region /rpt_family="GC_rich"
14801..14840
repeat_region /rpt_family="GC_rich"
complement(15362..15663)
repeat_region /rpt_family="AluJo"
15881..15924
repeat_region /rpt_family="GC_rich"
16110..16162
repeat_region /rpt_family="(CGGGG)n"
complement(16380..16658)
repeat_region /rpt_family="L3b"
complement(16694..16909)
repeat_region /rpt_family="L3"
17073..17108
repeat_region /rpt_family="GC_rich"
17379..17409
repeat_region /rpt_family="GC_rich"
18578..18710
repeat_region /rpt_family="MIR"
complement(18775..19088)
repeat_region /rpt_family="AluSx"
19780..19886
repeat_region /rpt_family="GA_rich"
19935..20031
repeat_region /rpt_family="MIR"
20334..20445
repeat_region /rpt_family="GA_rich"
20855..20960
repeat_region /rpt_family="(CTGGGG)n"
20906..21085
repeat_region /rpt_family="(TTGGGG)n"
21037..21216
repeat_region /rpt_family="(CTGGGG)n"
21166..21309
repeat_region /rpt_family="(TTGGGG)n"

repeat_region complement(22213..22416)
repeat_region /rpt_family="L2"
22417..22720
repeat_region /rpt_family="AluSg"
23332..23457
repeat_region /rpt_family="MIR"
complement(23477..23601)
repeat_region /rpt_family="MIR"
24030..24157

Query Match 24.7%; Score 644; DB 9; Length 134465;
Best Local Similarity 100.0%; Pred.No. 6.2e-136;
Matches 644; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 GAGTTCACAGGGCTGGATGAAGCCACTCTACTGCGGGCTCTGCAGGCCCTCAGCAGGAG 598
Db 116729 GAGTTCACAGGGCTGGATGAAGCCACTCTACTGCGGGCTCTGCAGGCCCTCAGCAGGAG 116788
QY 599 CACAAGGCCGAGATCATCATCTGTCTAGCGATGCCAGGCGTCAAGTTCTTCTAGCAGGGA 658
Db 116789 CACAAGGCCGAGATCATCATCTGTCTAGCGATGCCAGGCGTCAAGTTCTTCTAGCAGGGA 116848
QY 659 CTTGTCTCCCTTACTTCTTACCTCCACCTTTCAGGGCTTTCAAAAGAGACAGACC 718
Db 116849 CTTGTCTCCCTTACTTCTTACCTCCACCTTTCAGGGCTTTCAAAAGAGACAGACC 116908
QY 719 AGTGTCCCCCAAGACTGGATCTGTGACTCCACACAGACTCAAAAGGACTCCAGTCTCGAA 778
Db 116909 AGTGTCCCCCAAGACTGGATCTGTGACTCCACACAGACTCAAAAGGACTCCAGTCTCGAA 116968
QY 779 GGCTGGGACCTGGGGATGGGTTTCTCACACCCCATATGTCTGTCTCTTGGATAGGGTGAG 838
Db 116969 GGCTGGGACCTGGGGATGGGTTTCTCACACCCCATATGTCTGTCTCTTGGATAGGGTGAG 117028
QY 839 GCTGAAGCACACAGGAGAAAATATGTGCTTCTTCGCGCTTACCTCTTCCATCTCTAG 898
Db 117029 GCTGAAGCACACAGGAGAAAATATGTGCTTCTTCGCGCTTACCTCTTCCATCTCTAG 117088
QY 899 ACTGTCTTTGAGCCAGGCTCTGTAACCTGACACTTTATATGTGTTCACACATGTAAGTA 958
Db 117089 ACTGTCTTTGAGCCAGGCTCTGTAACCTGACACTTTATATGTGTTCACACATGTAAGTA 117148
QY 959 CATACACATGCGGCTGCAGACATGCTTCTGTCTCTCTCTCCCTCCACCCCTTTAGCT 1018
Db 117149 CATACACATGCGGCTGCAGACATGCTTCTGTCTCTCTCTCCCTCCACCCCTTTAGCT 117208
QY 1019 GCTGTTGCTCTCTCTCTCAGGCTGGTGTGATCTTCTTAGGGATGGGGAAAGCCCTG 1078
Db 117209 GCTGTTGCTCTCTCTCTCAGGCTGGTGTGATCTTCTTAGGGATGGGGAAAGCCCTG 117268
QY 1079 GCTGCAGGCAGCTTCCAGGCAATATGAAGATAGGAGGCCACAGGCTCTGGCAGTGAGAG 1138
Db 117269 GCTGCAGGCAGCTTCCAGGCAATATGAAGATAGGAGGCCACAGGCTCTGGCAGTGAGAG 117328
QY 1139 GTGTGGCCCCACACCGATTTATGATTTAAATCTCAACTCCCA 1182
Db 117329 GTGTGGCCCCACACCGATTTATGATTTAAATCTCAACTCCCA 117372

RESULT 11
AC016889
LOCUS AC016889 187723 bp DNA linear PRI 01-JAN-2003
DEFINITION Homo sapiens chromosome 17, clone RP11-506G7, complete sequence.
ACCESSION AC016889
VERSION AC016889.28 GI:27452920
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 187723)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 17, clone RP11-506G7

JOURNAL
REFERENCE
AUTHORS

Unpublished
2 (bases 1 to 187723)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Castle,A., Collangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Direct Submission
Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 187723)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J.,
Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Direct Submission
Submitted (24-NOV-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 187723)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J.,
Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (01-JAN-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jan 1, 2003 this sequence version replaced gi:25229251.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE
JOURNAL
COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information -----

Center project name: L5105
Center clone name: 506_G_7

Only the last 187.8 kilobases of this clone are being submitted.
The remainder overlaps accession number AC100793 [WICGR project
L20710].

FEATURES
source

FEATURES	source	Location/Qualifiers
repeat_region		1..187723 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="17" /map="17" /clone="RP11-506G7" /clone_lib="RPC1-11 Human Male BAC"
repeat_region		89..321 /rpt_family="AluJb"
repeat_region		322..337 /rpt_family="GA-rich"
repeat_region		338..648 /rpt_family="AluSx"
repeat_region		649..685 /rpt_family="GA-rich"
repeat_region		769..937 /rpt_family="AluSx"
repeat_region		1272..1331 /rpt_family="L1ME1"
repeat_region		1332..1628 /rpt_family="AluSg"
repeat_region		1629..2166 /rpt_family="L1ME1"
repeat_region		2179..2477 /rpt_family="AluSx"
repeat_region		complement(2615..2911) /rpt_family="AluSx"
repeat_region		2945..3253 /rpt_family="AluJb"
repeat_region		3521..3823 /rpt_family="AluY"
repeat_region		4035..4092 /rpt_family="T-rich"
repeat_region		complement(4111..4146) /rpt_family="Alu"
repeat_region		complement(4149..4235) /rpt_family="AluSc"
repeat_region		complement(4236..4377) /rpt_family="AluSg/x"
repeat_region		complement(4382..4681) /rpt_family="AluSc"
repeat_region		5013..5033 /rpt_family="AT-rich"
repeat_region		5076..5363 /rpt_family="AluSg"
repeat_region		complement(5415..5621) /rpt_family="AluY"
repeat_region		complement(5676..5869) /rpt_family="AluJo"
repeat_region		complement(5870..6161) /rpt_family="AluSc"
repeat_region		complement(6162..6273) /rpt_family="AluJo"
repeat_region		complement(6859..7124) /rpt_family="AluSx"
repeat_region		complement(7317..7614) /rpt_family="AluJb"
repeat_region		complement(7768..8068) /rpt_family="AluSg"
repeat_region		complement(8250..8413) /rpt_family="L2"
repeat_region		complement(8498..8794) /rpt_family="AluSx"
repeat_region		complement(8798..9096)

```
repeat_region      /rpt family="AluSx"
complement(9097..9390)
/rpt family="AluJo"
complement(10021..10164)
repeat_region      /rpt family="MER96"
10676..10718
/rpt family="GC_rich"
11098..11286
/rpt family="L1MC4a"
complement(11287..11570)
repeat_region      /rpt family="AluSx"
complement(11575..11702)
/rpt family="FLAM_C"
11703..11797
/rpt family="L1MC4a"
complement(11798..12106)
repeat_region      /rpt family="AluSq"
12107..12479
/rpt family="L1MC4a"
complement(12507..12739)
repeat_region      /rpt family="MIR"
complement(12855..12897)
repeat_region      /rpt family="MER5B"
complement(12898..13193)
/rpt family="AluSp"
complement(13194..13329)
/rpt family="MER5B"
complement(13393..13523)
repeat_region      /rpt family="L2"
13803..14115
/rpt family="AluJb"
complement(15989..16059)
/rpt family="MIR"
complement(16060..16357)
repeat_region      /rpt family="AluY"
complement(16358..16521)
/rpt family="MIR"
complement(17068..17245)
/rpt family="MIR"
complement(17358..17511)
/rpt family="MIR"
complement(17512..17804)
repeat_region

Query Match      24.7%; Score 644; DB 9; Length 187723;
Best Local Similarity 100.0%; Pred. No. 6.5e-136;
Matches 644; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 GAGTTCCACGGCTGGATGAAGCCACTCTACTGGGGGCTCTGACGGCCCTACAGCAGAG 598
Db 44582 GAGTTCCACGGCTGGATGAAGCCACTCTACTGGGGGCTCTGACGGCCCTACAGCAGAG 44641

QY 599 CACAAGGCGGAGATCATCACTGTGACGATGCGCCGAGGCGTCAAGTTCTTTAGCAGGGA 658
Db 44642 CACAAGGCGGAGATCATCACTGTGACGATGCGCCGAGGCGTCAAGTTCTTTAGCAGGGA 44701

QY 659 CTTGTCTCCCTTTACTTCTTACTCCCACTTCCAGGCTTTCAAAGGAGACAGACCC 718
Db 44702 CTTGTCTCCCTTTACTTCTTACTCCCACTTCCAGGCTTTCAAAGGAGACAGACCC 44761

QY 719 AGTGTCCCAAGACTGGATCTGTGACTTCCAACAGACTCAAAAGGACTCCAGTCCCTGAA 778
Db 44762 AGTGTCCCAAGACTGGATCTGTGACTTCCAACAGACTCAAAAGGACTCCAGTCCCTGAA 44821

QY 779 GCCTGGGACCTGGGATGGGTTTCTCACACCCCATATGTCTGTCCCTTGGATAGGTTGAG 838
Db 44822 GCCTGGGACCTGGGATGGGTTTCTCACACCCCATATGTCTGTCCCTTGGATAGGTTGAG 44881

QY 839 GCTGAAGCACAGGAGAGAAAATATGTGCTTCTTCGCCCTACCTCTTCCATCCCTAG 898
Db 44882 GCTGAAGCACAGGAGAGAAAATATGTGCTTCTTCGCCCTACCTCTTCCATCCCTAG 44941

QY 899 ACTGTCTTGAGCCAGGCTGTGTAAACCTTGACATTTTATATGTGTGTCAACATGTAAGTA 958
|||||
```

```
Db 44942 ACTGTCTTGAGCCAGGCTGTGTAAACCTTGACATTTTATATGTGTTCACATGTAAGTA 45001
QY 959 CATACACATGCGGCTGCAGACATGCTTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1018
Db 45002 CATACACATGCGGCTGCAGACATGCTTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 45061
QY 1019 GCTGTGCTCCCTTCTCAGGCTGTGTGATCTCTTCTAGGGATGGGGAGGCCCTG 1078
Db 45062 GCTGTGCTCCCTTCTCAGGCTGTGTGATCTCTTCTAGGGATGGGGAGGCCCTG 45121
QY 1079 GCTGAGGAGCCTTCCAGGCAATATGAAGATAGAGGCCCAACGGGCCCTGGCAGTGAGAG 1138
Db 45122 GCTGAGGAGCCTTCCAGGCAATATGAAGATAGAGGCCCAACGGGCCCTGGCAGTGAGAG 45181
QY 1139 GTGTGGCCCCACACGATTTATGATTTAAATCTCAACTCCCA 1182
Db 45182 GTGTGGCCCCACACGATTTATGATTTAAATCTCAACTCCCA 45225

RESULT 12
E50584
LOCUS E50584 531 bp DNA linear PAT 27-AUG-2002
DEFINITION Human protein and cDNA [3].
ACCESSION E50584
VERSION E50584.1 GI:22555007
KEYWORDS JP 2001161368-A/4.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 531)
AUTHORS Kato,S. and Saeiki,M.
TITLE Human protein and cDNA [3]
JOURNAL Patent: JP 2001161368-A 4 19-JUN-2001;
SCIENCE & TECH AGENCY
COMMENT OS Homo sapiens (human)
PN JP 2001161368-A/4
PD 19-JUN-2001
PF 06-DEC-1999 JP 1999346864
PI SEIJI KATO,MIHORO SAEKI
PC
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC
10//
PC C12P21/02,C12P21/08,C12N15/00,C12N5/00
CC
FH Key Location/Qualifiers.
FEATURES
source
1..531
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 20.4%; Score 531; DB 6; Length 531;
Best Local Similarity 100.0%; Pred. No. 1.7e-110;
Matches 531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 ATGGCGATGAGTTTCGAGTGGCCGCTGGCAGTATCGCTTCCACCCCTTTCTTACGTACAA 181
Db 1 ATGGCGATGAGTTTCGAGTGGCCGCTGGCAGTATCGCTTCCACCCCTTTCTTACGTACAA 60
QY 182 CCGAATGTGGACACTCGGCAGAGCAGCTGGCCGCTGGTGTCTGCTGGTCTGTCTGTCTTTC 241
Db 61 CCGAATGTGGACACTCGGCAGAGCAGCTGGCCGCTGGTGTCTGCTGTCTGTCTTTC 120
QY 242 TCCCGCTTCGCAAAACACTCCAGCATGACGTTGATGGAAGCTCAGGAGAGCCGCTCTTTC 301
Db 121 TCCCGCTTCGCAAAACACTCCAGCATGACGTTGATGGAAGCTCAGGAGAGCCGCTCTTTC 180
QY 302 AACACGTCRAGCTACAGGAAAGCTTCTCTGGAGTCGATCCAGATTGTATTAGAGGAA 361
Db 181 AACACGTCRAGCTACAGGAAAGCTTCTCTGGAGTCGATCCAGATTGTATTAGAGGAA 240
|||||
```

```
QY 362 CTGAGCAAGAAAGGAAACCTCGAGTGGTGGATGAAGCAAGTCCAGCTTCCTGATCATG 421
DB 241 CTGAGCAAGAAAGGAAACCTCGAGTGGTGGATGAAGCAAGTCCAGCTTCCTGATCATG 300
QY 422 TGGCGGAGGCGCAGAAAGTGGGGAAATCATCTATCAGTGGGTTCACAGGAGTGGCCAG 481
DB 301 TGGCGGAGGCGCAGAAAGTGGGGAAATCATCTATCAGTGGGTTCACAGGAGTGGCCAG 360
QY 482 AACAACTCCGTCTTTACCTGTATGAATGACTTAATGGGAAAGACACAGAGGATGAGGAG 541
DB 361 AACAACTCCGTCTTTACCTGTATGAATGACTTAATGGGAAAGACACAGAGGATGAGGAG 420
QY 542 TTCCAGGCGCTGGATGAAGCACTTACTGCGGGCTCTCAGGCCCTCAGCAGGAGCAC 601
DB 421 TTCCAGGCGCTGGATGAAGCACTTACTGCGGGCTCTCAGGCCCTCAGCAGGAGCAC 480
QY 602 AAGGCCGAGATCATCACTGTACAGCGATGCCGAGGCGTCAAGTTCCTCTAG 652
DB 481 AAGGCCGAGATCATCACTGTACAGCGATGCCGAGGCGTCAAGTTCCTCTAG 531

RESULT 13
AB014763
LOCUS Homo sapiens mRNA for DERP9 (dermal papilla derived protein 9),
DEFINITION complete cds.
ACCESSION AB014763
VERSION AB014763.1 GI:19909522
KEYWORDS DERP9 (dermal papilla derived protein 9).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Ikeda,A., Ukai,Y., Yamashita,M. and Yoshimoto,M.
TITLE Molecular cloning of a dermal papilla derived gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 531)
AUTHORS Ikeda,A., Ukai,Y., Yamashita,M. and Yoshimoto,M.
TITLE Direct Submission
JOURNAL Submitted (21-MAY-1998) Akiko Ikeda, Taisho Pharmaceutical Co.,
Ltd., Molecular Biology Lab.; 1-403 Yoshino-cho, Ohmiya, Saitama
330-8530, Japan (E-mail:s129i00ccm.taisho.co.jp, Tel:048-663-1111,
Fax:048-652-7254)
FEATURES
source
1..531
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
gene
1..531
/genes="derp9"
CDS
1..531
/genes="derp9"
/notes="similar to hypothetical 23.6 kd protein yuhl-ura8
intergenic region"
/codon_start=1
/product="DERP9 (dermal papilla derived protein 9)"
/protein_id="BAB87804.1"
/db_xref="GI:19909523"
/translation="MAMSEFWQYRFPFPFPLQNVDRQKLAACSLVLSFRLH
KQSMVMEASQPLFNIVKQLPVRISQIVLBELRKKGNLEWLDKSKSFLIMWR
RPEEWKLIYQWSSGQNNVSFTLYELTNGEDTDEBPHGLDEATLLRALQALQOEH
KABIIIVSDGRGVKPF"

ORIGIN
Query Match 20.4%; Score 531; DB 9; Length 531;
Best Local Similarity 100.0%; Pred. No. 1.7e-110;
Matches 531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 ATGGCGATGAGTTTCGAGTGGCGGTGGCAGTATCGCTCCACCCCTCTTTACGTTACAA 181
DB 1 ATGGCGATGAGTTTCGAGTGGCGGTGGCAGTATCGCTCCACCCCTCTTTACGTTACAA 60

US-08-731-499-4.rge
QY 182 CCGAATGTGGACATCGCGCAGAACGCTGGCGCCCTCGTGTGCTGTGCTGTGCTGTGCTTC 241
DB 61 CCGAATGTGGACATCGCGCAGAACGCTGGCGCCCTCGTGTGCTGTGCTGTGCTGTGCTTC 120
QY 242 TCGCGCCTGCACAAACAGTCCAGCATGACGGTGTGATGGAAGCTCAGGAGAGCCGCTCTTC 301
DB 121 TCGCGCCTGCACAAACAGTCCAGCATGACGGTGTGATGGAAGCTCAGGAGAGCCGCTCTTC 180
QY 302 AACAACTCCGTCTTTACCTGTATGAATGACTTAATGGGAAAGACACAGAGGATGAGGAA 361
DB 181 AACAACTCCGTCTTTACCTGTATGAATGACTTAATGGGAAAGACACAGAGGATGAGGAA 240
QY 362 CTGAGCAAGAAAGGAAACCTCGAGTGGTGGATGAAGCAAGTCCAGCTTCCTGATCATG 421
DB 241 CTGAGCAAGAAAGGAAACCTCGAGTGGTGGATGAAGCAAGTCCAGCTTCCTGATCATG 300
QY 422 TGGCGGAGGCGCAGAAAGTGGGGAAATCATCTATCAGTGGGTTCACAGGAGTGGCCAG 481
DB 301 TGGCGGAGGCGCAGAAAGTGGGGAAATCATCTATCAGTGGGTTCACAGGAGTGGCCAG 360
QY 482 AACAACTCCGTCTTTACCTGTATGAATGACTTAATGGGAAAGACACAGAGGATGAGGAG 541
DB 361 AACAACTCCGTCTTTACCTGTATGAATGACTTAATGGGAAAGACACAGAGGATGAGGAG 420
QY 542 TTCCAGGCGCTGGATGAAGCACTTACTGCGGGCTCTCAGGCCCTCAGCAGGAGCAC 601
DB 421 TTCCAGGCGCTGGATGAAGCACTTACTGCGGGCTCTCAGGCCCTCAGCAGGAGCAC 480
QY 602 AAGGCCGAGATCATCACTGTACAGCGATGCCGAGGCGTCAAGTTCCTCTAG 652
DB 481 AAGGCCGAGATCATCACTGTACAGCGATGCCGAGGCGTCAAGTTCCTCTAG 531

RESULT 14
AR412671
LOCUS Sequence 308 from patent US 6639063.
DEFINITION AR412671
ACCESSION AR412671
VERSION AR412671.1 GI:40167781
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
TITLE EST's and encoded human proteins
JOURNAL Patent: US 6639063-A 308 28-OCT-2003;
FEATURES Location/Qualifiers
source
1..495
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 18.3%; Score 476; DB 6; Length 495;
Best Local Similarity 99.6%; Pred. No. 6.4e-98;
Matches 476; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 109 CTGGGCTACTACGATGGCGATGAGTTTCAGTGGCGGTGGCAGTATCGCTTCCACCCCTT 168
DB 15 CTGGGCTACTACGATGGCGATGAGTTTCAGTGGCGGTGGCAGTATCGCTTCCACCCCTT 74
QY 169 CTTTACGTTTAAACCGAATGTGACACTCGGCAGAACGAGCTGGCGCGCTGTGCTCGCT 228
DB 75 CTTTACGTTTAAACCGAATGTGACACTCGGCAGAACGAGCTGGCGCGCTGTGCTCGCT 134
QY 229 GGTCTCTCTCTCTCTCGCGCTGCACAAACAGTCCAGCATGACGGTATGGAAGCTCAGGA 288
DB 135 GGTCTCTCTCTCTCTCGCGCTGCACAAACAGTCCAGCATGACGGTATGGAAGCTCAGGA 194
QY 289 GAGCCCGCTCTTCAACAAACAGTCAAGCTCAGCGAAAGCTTCTGTGGAGTCCGATCCAGAT 348
DB 195 GAGCCCGCTCTTCAACAAACAGTCAAGCTCAGCGAAAGSTTCTGTGGAGTCCGATCCAGAT 254
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 14:28:52 ; Search time 911.543 Seconds
(without alignments)
16917.381 Million cell updates/sec

Title: US-08-731-499-4
Perfect score: 2605
Sequence: 1 CAACTCGAATTAACCTC.....CTGGCGTTACCACTTAAT 2605

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2604.6	100.0	2605	AAV04699	Aav04699 Homo sapi
2	1089	41.8	1089	AAA40516	Aaa40516 Murine ad
3	1076.2	41.3	1243	AAI59037	Aai59037 Human pol
4	1076.2	41.3	1243	ADQ99260	Adq99260 DNA encod
5	1076.2	41.3	1243	ADB49020	Adb49020 Novel hum
6	1075	41.3	1098	AAI60823	Aai60823 Human pol
7	1073	41.2	1087	AAH68547	Aah68547 Human pol
8	1073	41.2	1100	ADQ84631	Adq84631 Human tum
9	603.6	23.2	634	ABQ57032	Abq57032 Human col
10	515.8	19.8	559	ABQ56941	Abq56941 Human col
11	489.2	18.8	675	ADQ50360	Adq50360 Novel can
12	464.4	17.8	507	ABQ56838	Abq56838 Human col
13	395	15.2	609	ABQ56915	Abq56915 Human col
14	353.6	13.6	545	ABX17514	Abx17514 Human CDN
15	317.4	12.2	349	ABQ56970	Abq56970 Human col
16	269	10.3	269	ABV89089	Abv89089 Human col
17	269	10.3	269	ABV88953	Abv88953 Human col
18	269	10.3	269	ABK39342	Abk39342 DNA encod
19	269	10.3	269	ACA11671	Aca11671 Human lung
20	269	10.3	269	ACA02857	Aca02857 Lung canc

C 21	269	10.3	269	10	ADH46899	Adh46899 Human lun
C 22	269	10.3	269	13	ADJ20818	Adj20818 Human lun
C 23	260	10.0	260	6	ABV86907	Abv86907 Human col
C 24	253.6	9.7	490	10	ADB56763	Adb56763 Toxicity-
25	253	9.7	304	3	AAA43467	Aaa43467 Mouse sec
26	238	9.1	276	2	AAV26672	Aav26672 Mouse nov
27	227	8.7	241	3	AAZ42852	Aaz42852 Human 5'
C 28	219.2	8.4	469	5	ABV37980	Abv37980 Human pro
C 29	218.2	8.4	335	5	ABV25168	Abv25168 Human pro
30	190	7.3	195	3	AAC00435	Aac00435 Human sec
C 31	187	7.2	409	5	AAF66957	Aaf66957 Novel hum
32	177.8	6.8	270	12	ADP28751	Adp28751 Human sec
C 33	169.4	6.5	217	4	AAI84862	Aai84862 Human pol
C 34	169.4	6.5	217	9	ACH31848	Ach31848 Human bon
35	166.6	6.4	246	2	AAV26673	Aav26673 Mouse nov
36	163.6	6.3	230	8	ABX38439	Abx38439 Bovine ES
C 37	156.4	6.0	1763	2	AAZ30709	Aaz30709 Rat neuro
C 38	156.4	6.0	1827	2	AAZ28264	Aaz28264 Rat neuro
C 39	154.2	5.9	2201	4	AAZ26054	Aaz26054 Human CDN
C 40	154.2	5.9	2201	8	ABX73395	Abx73395 Human nov
41	152.6	5.9	3315	5	ABA14281	Abai4281 Human ner
C 42	146.2	5.6	8710	2	AAV32370	Aav32370 Complete
C 43	146.2	5.6	8710	3	AAAS9039	Aas9039 Nucleotid
C 44	146.2	5.6	8710	6	ABA94248	Abas94248 Adenoviru
C 45	146.2	5.6	8710	10	ADB75145	Adb75145 Plaemid p

ALIGNMENTS

RESULT 1
AAV04699
ID AAV04699 standard; CDNA; 2605 BP.
XX
AC AAV04699;
XX
DT 21-JUL-1998 (first entry)
XX
DE Homo sapiens 20q13 amplicon cc43 transcript.

XX 20q13 amplicon; chromosome 20; tumour; detection;
KW chromosomal abnormalities; probe; gene therapy; antisense inhibition;
KW treatment; age-related macular degeneration; retinitis pigmentata;
KW Leber's congenital amaurosis; ds.

OS Homo sapiens.
XX
PN WO9802539-A1.
XX
PD 22-JAN-1998.
XX
PP 15-JUL-1997; 97WO-US012343.

XX 15-JUL-1996; 96US-00680395.
PR 16-OCT-1996; 96US-00731499.
PR 17-JAN-1997; 97US-00785532.
XX
PA (REGC) UNIV CALIFORNIA.

XX Gray JW, Collins CC, Hwang S, Godfrey T, Kowbel D, Rommens J;
WPI; 1998-110587/10.

XX New sequences from the 20q13 amplicon - used for detecting chromosomal
PT abnormalities, particularly tumours, and for developing products for
treating diseases.

XX Claim 1; Page 62; 91pp; English.

XX The sequence is that of a cDNA sequence cc43, which was isolated from the
CC 20q13 amplicon. It is expressed in normal tissues but not been found in
CC the breast cancer cell line. It can be used as a probe for the detection
CC of chromosomal abnormalities at 20q13. It and other sequences isolated

QY 1921 CTTCTCTTAAGAATTCATATTCCTTAATTCCTTTATCTGGGTAACTAGATTTTATCAACA 1980
DB 1921 CTTCTCTTAAGAATTCATATTCCTTAATTCCTTTATCTGGGTAACTAGATTTTATCAACA 1980
QY 1981 CTTTCAGATCTGCGGTAAACCTCTCTTATAGAACCTGTGATGACACTGCTCTCTTC 2040
DB 1981 CTTTCAGATCTGCGGTAAACCTCTCTTATAGAACCTGTGATGACACTGCTCTCTTC 2040
QY 2041 TCCAAACATCTCACACGACACATGTAGACTAGATTAGAACCTCTCTGTTTTCTTTTCA 2100
DB 2041 TCCAAACATCTCACACGACACATGTAGACTAGATTAGAACCTCTCTGTTTTCTTTTCA 2100
QY 2101 TACTTTTCTATCATGCTTCCCTCAATTAATAATTTTTATGTGTGTGAATGCTG 2160
DB 2101 TACTTTTCTATCATGCTTCCCTCAATTAATAATTTTTATGTGTGTGAATGCTG 2160
QY 2161 CCCCAAGTCAGTTTCTCTCACTAACTATAAATCCGTAAGCTGGGATCCTTCCAAATTT 2220
DB 2161 CCCCAAGTCAGTTTCTCTCACTAACTATAAATCCGTAAGCTGGGATCCTTCCAAATTT 2220
QY 2221 GATCACCCTTAGTACAGTAGGAACACAGTAAGATTCAATTTGTTGGAATGAAT 2280
DB 2221 GATCACCCTTAGTACAGTAGGAACACAGTAAGATTCAATTTGTTGGAATGAAT 2280
QY 2281 GAATGAATTTGTTTGTAGTAAAGTCTGGGGAAACCCAGGTGAGAAAGCCTAGAAAGCA 2340
DB 2281 GAATGAATTTGTTTGTAGTAAAGTCTGGGGAAACCCAGGTGAGAAAGCCTAGAAAGCA 2340
QY 2341 GGTGGAATCCAGGCTAGATAGACTTAGTGTACTCAAGAAAGGCTAGCTGAATATAA 2400
DB 2341 GGTGGAATCCAGGCTAGATAGACTTAGTGTACTCAAGAAAGGCTAGCTGAATATAA 2400
QY 2401 GGTTCAAATTTAGTCAAGAATAGTCAAGACATGGGCAAGAGTGTCTGCTGCTGCC 2460
DB 2401 GGTTCAAATTTAGTCAAGAATAGTCAAGACATGGGCAAGAGTGTCTGCTGCTGCC 2460
QY 2461 GAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCCGTACCAATTC 2520
DB 2461 GAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCCGTACCAATTC 2520
QY 2521 GGCCTATAGTGTGCTGATTTACAAATTCACCTGGCGGTCTTTTACACGTCGTGACTGGGA 2580
DB 2521 GGCCTATAGTGTGCTGATTTACAAATTCACCTGGCGGTCTTTTACACGTCGTGACTGGGA 2580
QY 2581 AAACCCCTGGCTTACCCCAACTTAAT 2605
DB 2581 AAACCCCTGGCTTACCCCAACTTAAT 2605

RESULT 2
ID AAA40516
AC AAA40516;
XX
XX
DT 16-NOV-2000 (first entry)
DE Murine adult spleen cDNA fragment AE402_li.
XX
XX Secreted protein; cytostatic; immunostimulatory; antimicrobial;
KW antiviral; immunosuppressive; antiinflammatory; vulnerary; cytokine;
KW cell proliferation; differentiation; regulator; treatment; tumor;
KW autoimmune disease; inflammatory disorder; wound; microbial infection;
KW viral disease; graft versus host reaction suppression; ss.
XX
OS Mus sp.
XX
PN WO200037630-A1.
XX
PD 29-JUN-2000.
XX
PF 22-DEC-1999; 99WO-US031005.
XX

PR 23-DEC-1998; 98US-00220876.
XX (GEM) GENETICS INST INC.
XX
PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Bowman MR;
XX WPI; 2000-442661/38.
DR P-PSDB; AAB10244.
XX
XX Secreted human proteins AS296-li and AS34-li, useful for treating tumors,
PT autoimmune diseases, inflammatory disorders, wounds, microbial infections
PT and viral diseases.
XX
PS Disclosure; Page 225-226; 293pp; English.
XX
XX This invention describes novel secreted human proteins (I) which have
CC cytostatic, immunostimulatory, antimicrobial, antiviral,
CC immunosuppressive, antiinflammatory and vulnerary activity and which act
CC as cytokine, cell proliferation or differentiation regulators. (I) is
CC useful for treating tumors, autoimmune diseases, inflammatory disorders,
CC wounds, microbial infections and viral diseases. (I) is also useful for
CC suppressing graft versus host reaction. AAA40490-A40580 represent cDNA
CC fragments that encode the secreted proteins AAB10226-B10288 described in
CC the method of the invention
XX
SQ Sequence 1089 BP; 263 A; 299 C; 284 G; 243 T; 0 U; 0 Other;
Query Match 41.8%; Score 1089; DB 3; Length 1089;
Best Local Similarity 100.0%; Pred. No. 6.7e-214;
Matches 1089; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 111 GGGCTACTACGATGGCGATGAGTTTCGAGTGGCGGTGGCAGTATCGCTTCCACCCCTTCT 170
DB 1 GGGCTACTACGATGGCGATGAGTTTCGAGTGGCGGTGGCAGTATCGCTTCCACCCCTTCT 60
QY 171 TTACGTTCACACCGAATGTGGACATCGGCAGACAGCAGTGGCGGCTGGTGTCTGCTGG 230
DB 61 TTACGTTCACACCGAATGTGGACATCGGCAGACAGCAGTGGCGGCTGGTGTCTGCTGG 120
QY 231 TCCTGTCTCTTCGCGCCTGCACAAACAGTCAGCATGACGGTGATGGAGCTCAGAGA 290
DB 121 TCCTGTCTCTTCGCGCCTGCACAAACAGTCAGCATGACGGTGATGGAGCTCAGAGA 180
QY 291 GCCCGCTCTTCAACACAGTCAAGCTACAGCAAGAGCTTCTCTGTGGAGTCGATCCAGATTG 350
DB 181 GCCCGCTCTTCAACACAGTCAAGCTACAGCAAGAGCTTCTCTGTGGAGTCGATCCAGATTG 240
QY 351 TATTAGAGAACTGAGGAAAGAAAGGAACTCGAGTGGTTGGATTAAGAGCAAGTCCAGCT 410
DB 241 TATTAGAGAACTGAGGAAAGAAAGGAACTCGAGTGGTTGGATTAAGAGCAAGTCCAGCT 300
QY 411 TCCTGATCATGTGGCGGAGCCAGAGAAATGGGGAACCTCATCTATCAGTGGTTCCTCA 470
DB 301 TCCTGATCATGTGGCGGAGCCAGAGAAATGGGGAACCTCATCTATCAGTGGTTCCTCA 360
QY 471 GGAGTGGCCAGAAACAACTCCGTCTTTTACCTGTATGAACTGACTAATGGGGAAGACACAG 530
DB 361 GGAGTGGCCAGAAACAACTCCGTCTTTTACCTGTATGAACTGACTAATGGGGAAGACACAG 420
QY 531 AGGATGAGAGTTCACCGGCTGGATGAAGCACTTACTTGGGGGTCTGTCAGGCCCTTAC 590
DB 421 AGGATGAGAGTTCACCGGCTGGATGAAGCACTTACTTGGGGGTCTGTCAGGCCCTTAC 480
QY 591 AGCAGGAGCACAAGCCCGAGATCATCATGTGATGGATGCGGAGGCGTCAAGTCTTCT 650
DB 481 AGCAGGAGCACAAGCCCGAGATCATCATGTGATGGATGCGGAGGCGTCAAGTCTTCT 540
QY 651 AGCAGGAGCCTGTCTCCCTTTTACTTCTTACCTCCCACTTTTCCAGGGCTTTTCAAAAGGAG 710
DB 541 AGCAGGAGCCTGTCTCCCTTTTACTTCTTACCTCCCACTTTTCCAGGGCTTTTCAAAAGGAG 600
QY 711 ACAGACCCAGTGTCCCCCAAAGACTGTGTGACTCCACAGACTCAAAAGGACTCCA 770

Db 601 ACAGACCCAGTGTCCCCAAAGACTGGATCTGTGACTCCACCAGACTCAAAAGGACTCCA 660
Qy 771 GTCCTGAAGCTGGGACCTGGGATGGTTTCTCACACCCATATGTCTGCTCCCTTGGAT 830
Db 661 GTCCTGAAGCTGGGACCTGGGATGGTTTCTCACACCCATATGTCTGCTCCCTTGGAT 720
Qy 831 AGGGTGAGCTGAAGCACCAGGAGAAAATATGTCTTCTTCGCCCCCTACCTCTTCC 890
Db 721 AGGGTGAGCTGAAGCACCAGGAGAAAATATGTCTTCTTCGCCCCCTACCTCTTCC 780
Qy 891 CATCTAGACTGTCTTGAGCCAGGCTCTGTAACCTGCACCTTTATATGTGTTCCACACA 950
Db 781 CATCTAGACTGTCTTGAGCCAGGCTCTGTAACCTGCACCTTTATATGTGTTCCACACA 840
Qy 951 TGTAAGTACATACACATGCGCTGACGACATGCTTCTGCTCTCTCCCTCCACGCC 1010
Db 841 TGTAAGTACATACACATGCGCTGACGACATGCTTCTGCTCTCTCCCTCCACGCC 900
Qy 1011 CTTTAGCTGTCTTCCCTCTCTCAGCTGCTGCTGATCCTTCTAGGGATGGGG 1070
Db 901 CTTTAGCTGTCTTCCCTCTCTCAGCTGCTGCTGATCCTTCTAGGGATGGGG 960
Qy 1071 AAGCCCTGGCTGACGAGCCTTCCAGGCAATATGAAGATAGGAGCCACCGGCTTGGC 1130
Db 961 AAGCCCTGGCTGACGAGCCTTCCAGGCAATATGAAGATAGGAGCCACCGGCTTGGC 1020
Qy 1131 AGTGAGAGTGTGGCCCCACACCGATTTATGATATTAATCTCAACTCCCAAAAAA 1190
Db 1021 AGTGAGAGTGTGGCCCCACACCGATTTATGATATTAATCTCAACTCCCAAAAAA 1080
Qy 1191 AAAAAAAA 1199
Db 1081 AAAAAAAA 1089

RESULT 3

AA159037
ID AA159037 standard; cDNA; 1243 BP.
XX
AC AA159037;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1240.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HYSE-) HYSEQ INC.
XX

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
DR P-PSDB; AA39881.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
PS Claim 1; SEQ ID NO 1240; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AA38642-AA42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX

Sequence 1243 BP; 313 A; 321 C; 326 G; 283 T; 0 U; 0 Other;

Query Match 41.3%; Score 1076.2; DB 4; Length 1243;
Best Local Similarity 99.7%; Pred. No. 3e-211;
Matches 1078; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 113 GCTACTACGATGGCGATGAGTTTCGAGTGGCGGTGGAGTATCGTTCGCCACCTCTCTTT 172
Db 163 GCGTCGACGATGGCGATGAGTTTCGAGTGGCGGTGGAGTATCGTTCGCCACCTCTCTTT 222
Qy 173 ACCTTACAAACGATGTGGACACTCGGCAGAGCAGCTGGCGCTGTGCTGCTGCTGCTGCTC 232
Db 223 ACCTTACAAACGATGTGGACACTCGGCAGAGCAGCTGGCGCTGTGCTGCTGCTGCTGCTC 282
Qy 233 CTGTCCTTCTGCGCTGCACAAACAGTCCAGCATGACGGTGATGGAAGCTCAGGAGAGC 292
Db 283 CTGTCCTTCTGCGCTGCACAAACAGTCCAGCATGACGGTGATGGAAGCTCAGGAGAGC 342
Qy 293 CCGCTCTTCAACAGCTCAAGCTACAGGAAAGCTTCTGTGGAGTCAATCAGATTTGTA 352
Db 343 CCGCTCTTCAACAGCTCAAGCTACAGGAAAGCTTCTGTGGAGTCAATCAGATTTGTA 402
Qy 353 TTAGAGGAACCTGAGGAAGAAAGGAACTCGAGTGGTTGGATAAGAGCAAGTCCAGCTTC 412
Db 403 TTAGAGGAACCTGAGGAAGAAAGGAACTCGAGTGGTTGGATAAGAGCAAGTCCAGCTTC 462
Qy 413 CTGATCATGTGCGGAGGCCAGAAAGTGGGGAAACTCATCTATCAGTGGGTTTCCAGG 472
Db 463 CTGATCATGTGCGGAGGCCAGAAAGTGGGGAAACTCATCTATCAGTGGGTTTCCAGG 522
Qy 473 AGTGCCAGAAACAATCTCGTCTTTTACCTGTATGAACTGAATAATGGGGAAGACACAGAG 532
Db 523 AGTGCCAGAAACAATCTCGTCTTTTACCTGTATGAACTGAATAATGGGGAAGACACAGAG 582
Qy 533 GATGAGGATTTCCAGGGCTGGATGAAGCCACTCTACTCGCGGCTCTCAGGCCCTACAG 592
Db 583 GATGAGGATTTCCAGGGCTGGATGAAGCCACTCTACTCGCGGCTCTCAGGCCCTACAG 642
Qy 593 CAGGAGCAAAAGGCCGAGATCATCTGTTCAGCGATGGCCGAGGCGTCAAGTCTTCTTAG 652
Db 643 CAGGAGCAAAAGGCCGAGATCATCTGTTCAGCGATGGCCGAGGCGTCAAGTCTTCTTAG 702
Qy 653 CAGGAGCACTGTCTCTCTTTACTTCTTACTCTCCACCTTTCCAGGGCTTCAAAAGAGAC 712
Db 703 CAGGAGCACTGTCTCTCTTTACTTCTTACTCTCCACCTTTTCCAGGGCTTCAAAAGAGAC 762

QY 1073 GCCTGGCTGCAGGCGCTTCAGGCAATATGAGATAGGAGCCGCGCTGGCAG 1132
Db 1123 GCCTGGCTGCAGGCGCTTCAGGCAATATGAGATAGGAGCCGCGCTGGCAG 1182
QY 1133 TGAGAGGTGTGGCCGCCACACCGGATTTATGATATTAAATCTCAACTCCCAAAAAA 1192
Db 1183 TGAGAGGTGTGGCCGCCACACCGGATTTATGATATTAAATCTCAACTCCCAAAAAA 1242
QY 1193 A 1193
Db 1243 A 1243

RESULT 6

AAI60823/c
ID AAI60823 standard; cDNA; 1098 BP.
AC AAI60823;
XX
DT 22-OCT-2001 (first entry)
DE Human polynucleotide SEQ ID NO 4812.
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia; ss.
OS Homo sapiens.
XX
PN WO200153312-A1.
PD 26-JUL-2001.
PF 26-DEC-2000; 2000WO-US034263.
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
PA (HYSE-) HYSEQ INC.
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM41667.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.
PT
PS Claim 1; SEQ ID NO 4812; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAM42213) with nootropic, immunosuppressant and cytosstatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening, CC assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification
XX
SQ Sequence 1098 BP; 245 A; 287 C; 308 G; 258 T; 0 U; 0 Other;
Query Match 41.3%; Score 1075; DB 4; Length 1098;
Best Local Similarity 99.9%; Pred. No. 5,1e-211;
Matches 1086; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 108 GCTGGGCTACTAGGATGGCGATGAGTTTCGAGTGGCGCTGGCAGTATCGTTCACCCCT 167
Db 1087 GCTGGGCTACTAGGATGGCGATGAGTTTCGAGTGGCGCTGGCAGTATCGTTCACCCCT 1028
QY 168 TCTTTACGTTTCAACCGGAATGTGGACACTCGGCAGAAAGCAGCTGGCGCCTCGTGTCTGC 227
Db 1027 TCTTTACGTTTCAACCGGAATGTGGACACTCGGCAGAAAGCAGCTGGCGCCTCGTGTCTGC 968
QY 228 TGGTCCTGTCTTCTGCGCCTTGCAAAACAGTCCAGCATGACGGTGTGAAAGCTCAGG 287
Db 967 TGGTCCTGTCTTCTGCGCCTTGCAAAACAGTCCAGCATGACGGTGTGAAAGCTCAGG 908
QY 288 AGAGCCCGCTCTTCAACAGCTCAAGCTACAGCGAAAGCTTCTCTGGAGTCCATCCAGA 347
Db 907 AGAGCCCGCTCTTCAACAGCTCAAGCTACAGCGAAAGCTTCTCTGGAGTCCATCCAGA 848
QY 348 TTGTATTACAGAACTCGAGGAAGAAAGGAACTCGAGTGGTGGTAAAGCAGCAAGTCCA 407
Db 847 TTGTATTACAGAACTCGAGGAAGAAAGGAACTCGAGTGGTGGTAAAGCAGCAAGTCCA 788
QY 408 GCTTCCTGATCATGTGGCGGAGCCAGAAAGAAATGGGGGAAACTCATCTATCAGTGGGTTT 467
Db 787 GCTTCCTGATCATGTGGCGGAGCCAGAAAGAAATGGGGGAAACTCATCTATCAGTGGGTTT 728
QY 468 CCAGAGTGGCCAGAACTCGGCTTTTACCTGTATGAATGACTTAATGGGGAAGACA 527
Db 727 CCAGAGTGGCCAGAACTCGGCTTTTACCTGTATGAATGACTTAATGGGGAAGACA 668
QY 528 CAGAGGATGAGGATTTCCACGGCTGGATGAAGCCACTCTACTGCGGGCTCTCAGGCC 587
Db 667 CAGAGGATGAGGATTTCCACGGCTGGATGAAGCCACTCTACTGCGGGCTCTCAGGCC 608
QY 588 TACAGCAGGAGCAAGGCGCGAGATCATCACTGTTCAGCGATGG - CCGAGCGCTCAAGTTC 646
Db 607 TACAGCAGGAGCAAGGCGCGAGATCATCACTGTTCAGCGATGGCCCGAGGCGTCAAGTTC 548
QY 647 TTCTAGCAGGACCTGTCTCCCTTTTACTTCTTACCTTCCACCTTTCCAGGGCTTTCAAAA 706
Db 547 TTCTAGCAGGACCTGTCTCCCTTTTACTTCTTACCTTCCACCTTTCCAGGGCTTTCAAAA 488
QY 707 GGAGACAGACCCAGGTGTCCCAAGAGACTGGATCTGTGACTCCACAGACTCAAAAGGAC 766
Db 487 GGAGACAGACCCAGGTGTCCCAAGAGACTGGATCTGTGACTCCACAGACTCAAAAGGAC 428
QY 767 TCCAGTCTCTGAAGGCTGGGACCTGGGGATGGGTTTCTCACACCCCATATGTCTGTCCCTT 826
Db 427 TCCAGTCTCTGAAGGCTGGGACCTGGGGATGGGTTTCTCACACCCCATATGTCTGTCCCTT 368
QY 827 GGATAGGTGAGGCTGAAGCACACAGGAGAAATATGTGCTTCTCGCCCTACCTCCT 886
Db 367 GGATAGGTGAGGCTGAAGCACACAGGAGAAATATGTGCTTCTTCTCGCCCTACCTCCT 308
QY 887 TTCCCATCTAGACTGTCTCTTTCAGCCAGGCTGTGTAACCTGACACTTTATATGTGTCA 946
Db 307 TTCCCATCTAGACTGTCTCTTTCAGCCAGGCTGTGTAACCTGACACTTTATATGTGTCA 248
QY 947 CACATGTAGTACATACACATATGGCCCTGCAGCAGCATGCTTCTGTCTCTCTCTCC 1006
Db 247 CACATGTAGTACATACACATATGGCCCTGCAGCAGCATGCTTCTGTCTCTCTCTCTCC 188
QY 1007 ACCCTTTAGTGTGTGCTTCCCTTCTCAGGCTGGTGTGATCCTTCTCTAGGGGATG 1066

RESULT 8
ADQ84631
ID ADQ84631 standard; cDNA; 1100 BP.
XX
AC ADQ84631;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #1445.
XX
KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
KW cancer; cell proliferative disorder; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004060270-A2.
XX
XX 22-JUL-2004.
XX
PD 15-OCT-2003; 2003WO-US029126.
XX
XX 18-OCT-2002; 2002US-0418988P.
XX
XX (GETH) GENENTECH INC.
PA (WUTD/) WU T D.
PA (ZHOU/) ZHOU Y.
XX
XX Wu TD, Zhou Y;
PI
XX WPI; 2004-534300/51.
DR
XX
PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
PT preventing or treating cell proliferative disorders such as cancer.
XX
PS Claim 1; SEQ ID NO 1445; 5504pp; English.
XX
CC The present invention describes an isolated tumour-associated antigenic
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC a process for producing a polypeptide; (4) an isolated polypeptide
CC comprising: (a) an amino acid sequence encoded by any of the above
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC length coding region of the above nucleotide sequences; or (c) a sequence
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a

CC human TAT cDNA sequence from the present invention.
XX
SQ Sequence 1100 BP; 247 A; 308 C; 292 G; 253 T; 0 U; 0 Other;
Query Match 41.2%; Score 1073; DB 13; Length 1100;
Best Local Similarity 100.0%; Pred. No. 1.3e-210;
Matches 1073; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 109 CTGGGCTACTAGCATGGCGATGAGTTTCGAGTGGCGTGGCAGTATCGCTTCCACCCCTT 168
DB 28 CTGGGCTACTAGCATGGCGATGAGTTTCGAGTGGCGTGGCAGTATCGCTTCCACCCCTT 87
QY 169 CTTTACGTTTAAACCGAATGTGACACACTCGGAGAGAGAGCTGGCCCTGTGTCGCT 228
DB 88 CTTTACGTTTAAACCGAATGTGACACACTCGGAGAGAGAGCTGGCCCTGTGTCGCT 147
QY 229 GGTCTGTCTTCTGCGGCTGCACAAACAGTCCAGCATGACGGTATGGAAGCTCAGGA 288
DB 148 GGTCTGTCTTCTGCGGCTGCACAAACAGTCCAGCATGACGGTATGGAAGCTCAGGA 207
QY 289 GAGCCCGCTCTTCAACACAGCTCAAGCTACAGCGAAGCTTCTCTGTGGAGTCCGATCCAGAT 348
DB 208 GAGCCCGCTCTTCAACACAGCTCAAGCTACAGCGAAGCTTCTCTGTGGAGTCCGATCCAGAT 267
QY 349 TGTATTAGAGAACTGAGGAAGAAAGGAACTCTGAGTGGTGGATAGAGCAAGTCCAG 408
DB 268 TGTATTAGAGAACTGAGGAAGAAAGGAACTCTGAGTGGTGGATAGAGCAAGTCCAG 327
QY 409 CTTCTGTATCATGTGGCGGAGCCAGAGAAATGGGGGAACTCATCTATCAGTGGGTTTC 468
DB 328 CTTCTGTATCATGTGGCGGAGCCAGAGAAATGGGGGAACTCATCTATCAGTGGGTTTC 387
QY 469 CAGGAGTGCCAGAACTCCGCTCTTTACCTCTATGAACTGACTTAATGGGGAAGACAC 528
DB 388 CAGGAGTGCCAGAACTCCGCTCTTTACCTCTATGAACTGACTTAATGGGGAAGACAC 447
QY 529 AGAGGATGAGGAGTTCACAGGCTCGATGAAGACCACTTACTGCGGGCTCTGCAGGCCCT 588
DB 448 AGAGGATGAGGAGTTCACAGGCTCGATGAAGACCACTTACTGCGGGCTCTGCAGGCCCT 507
QY 589 ACAGCAGGAGCAGAGGCGGAGATCATCTCTGACGATGCGGAGGCGTCAAGTTCCT 648
DB 508 ACAGCAGGAGCAGAGGCGGAGATCATCTCTGACGATGCGGAGGCGTCAAGTTCCT 567
QY 649 CTAGCAGGACCTGCTCTCTTACTTCTTACTCTCCACCTTTCACAGGCTTTCACAAAGG 708
DB 568 CTAGCAGGACCTGCTCTCTTACTTCTTACTCTCCACCTTTCACAGGCTTTCACAAAGG 627
QY 709 AGACAGACCCAGTGTCCCCCAAAGACTGGATCTGACTTCCACAGACTCAAAAGGACTC 768
DB 628 AGACAGACCCAGTGTCCCCCAAAGACTGGATCTGACTTCCACAGACTCAAAAGGACTC 687
QY 769 CAGTCTGAGGCTGGGACCTGGGGATGGGTTTCTCACACCCCATATGTCCTCCCTGG 828
DB 688 CAGTCTGAGGCTGGGACCTGGGGATGGGTTTCTCACACCCCATATGTCCTCCCTGG 747
QY 829 ATAGGCTGAGGCTGAAGCAGCAGGAGAAATATGTGCTTCTTCTCGCCCTTACCTCTT 888
DB 748 ATAGGCTGAGGCTGAAGCAGCAGGAGAAATATGTGCTTCTTCTCGCCCTTACCTCTT 807
QY 889 CCCATCTAGACTGTCTTGTAGCCAGGGTCTGTAAACCTGACACTTATATGTGTTTACA 948
DB 808 CCCATCTAGACTGTCTTGTAGCCAGGGTCTGTAAACCTGACACTTATATGTGTTTACA 867
QY 949 CATGTAGTACATACACATGCGGCTGAGCAGACATGCTTCTGCTCTCTCTCTCCCTCC 1008
DB 868 CATGTAGTACATACACATGCGGCTGAGCAGACATGCTTCTGCTCTCTCTCTCTCCCTCC 927
QY 1009 CCCTTTAGTGTGCTGCTCTCTCTCTGAGGCTGTGCTGGATCTTCTCTCGGAGATGGG 1068
DB 928 CCCTTTAGTGTGCTGCTCTCTCTCTGAGGCTGTGCTGGATCTTCTCTCGGAGATGGG 987
QY 1069 GGAAGCCCTGGCTGCAGGCGAGCCTTTCAGGCAATATGAGATAGGAGGCCCGGCGCTG 1128

```

Db      988  GGAACCCCTGGCTGCAGGAGCCTTCCAGGCAATATAGATAGGAGGCCCAAGGCGCTG 1047
QY      1129  GCAGTGAGAGGTGTGGCCCCACACCGATTATGATATTAAATCTCAACTCCC 1181
Db      1048  GCAGTGAGAGGTGTGGCCCCACACCGATTATGATATTAAATCTCAACTCCC 1100

RESULT 9
ABQ57032/c
ID      ABQ57032 standard; cDNA; 634 BP.
XX      AC      ABQ57032;
XX      XX
XX      DT      02-AUG-2002 (first entry)
XX      DE      Human colon cancer related nucleotide sequence SEQ ID NO:727.
XX      KW      Human; colon cancer; cancer; tissue profiling; forensic; mapping;
XX      KW      genetic analysis; diagnostic; antisense therapy; gene; ss.
XX      OS      Homo sapiens.
XX      PN      WO200229086-A2.
XX      PD      11-APR-2002.
XX      PF      02-OCT-2001; 2001WO-US030732.
XX      PR      02-OCT-2000; 2000US-0237271P.
XX      PA      (FARB ) BAYER CORP.
XX      PI      Burgess C, Astle JH, Carroll E, Catino TU, Dwivedi P, Molino GA;
XX      PI      Thiaglingam A, Lewis MB;
XX      DR      WPI; 2002-426115/45.
XX      PT      New isolated nucleic acid that is differentially expressed in cancer
XX      PT      tissues useful for determining the presence of colon cancer in a cell or
XX      PT      tissue type, and in antisense therapy.
XX      PS      Claim 1; Fig 1; 796pp; English.
XX      CC      ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
XX      CC      expressed in cancer tissues. ABQ56306 to ABQ56306 represent proteins
XX      CC      encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
XX      CC      used in antisense therapy. An antibody immunoreactive with a polypeptide
XX      CC      encoded by (I) is useful for detecting cancer in a patient sample, and
XX      CC      for detecting the presence or absence of a polynucleotide encoded by a
XX      CC      nucleic acid which hybridizes to (I) in a cell. A probe/primer derived
XX      CC      from (I) can be used for determining the presence of a nucleic acid which
XX      CC      hybridizes to (I), and for determining the phenotype of cells in a sample
XX      CC      of cells from a patient. (I) is useful for determining the presence of
XX      CC      colon cancer in a cell or tissue type, for determining the presence or
XX      CC      state of other type of cancer, in antisense therapy, to generate
XX      CC      macroarrays on a solid surface, to identify a chromosome on which the
XX      CC      corresponding gene resides, and in tissue profiling, forensics, genetic
XX      CC      analysis, mapping and diagnostic applications. (I) can be used to raise
XX      CC      antibodies, and to screen for peptide analogues and antagonists
XX      SQ      Sequence 634 BP; 147 A; 168 C; 160 G; 158 T; 0 U; 1 Other;

Query Match      23.2%; Score 603.6; DB 6; Length 634;
Best Local Similarity 98.7%; Pred. No. 3.5e-114;
Matches 629; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

QY      321  GAAAGCTTCCTGTGAGTCGATCCAGATTGTATTAGAGCACTGAGGAAGGAAGCAACC 380
Db      634  GAAAGCTTCCTGTGAGTCGATCCAGATTGTATTAGAGCACTGAGGAAGGAAGCAACC 577
QY      381  TCGAGTGGTTGGATGAAGAGCAAGTCAGCTTCTCTGATCATGTGGCGGAGGCCAGAGAAT 440

```

```

Db      576  TCGAGTGGTTGGATGAAGAGCAAGTCCAGCTTCTCTGATCATGTGGCGGAGGCCAGAGAAT 517
QY      441  GGGGGAAACTCATCTATCAGTGGGTTTCCAGAGTGGCCAGAACAACTCCCGTCTTTACCC 500
Db      516  GGGGGAAACTCATCTATCAGTGGGTTTCCAGAGTGGCCAGAACAACTCCCGTCTTTACCC 458
QY      501  TGTATGAATGACTAATGGGGAAGACACAGAGGATGAGGAGTTCCACGGGCTGGATGAAG 560
Db      457  TGTATGAATGACTAATGGGGAAGACACAGAGGATGAGGAGTTCCACGGGCTGGATGAAG 398
QY      561  CCACTTACTCGGGCTCTGAGGCCCTTACAGCAGGAGCAAGGCCGAGATCATCATCTG 620
Db      397  CCACTTACTCGGGCTCTGAGGCCCTTACAGCAGGAGCAAGGCCGAGATCATCATCTG 338
QY      621  TCAGCGATGGCGAGGCGTCAAGTTCTTTAGCAGGAGCTGTCTCCCTTTACTTCTTAC 680
Db      337  TCAGCGATGGCGAGGCGTCAAGTTCTTTAGCAGGAGCTGTCTCCCTTTACTTCTTAC 278
QY      681  CTCCCACTTTCCAGGGCTTTCAAAAGGAGACAGACCCAGTGTCTCCCCCAAAGACTGGATC 740
Db      277  CTCCCACTTTCCAGGGCTTTCAAAAGGAGACAGACCCAGTGTCTCCCCCAAAGACTGGATC 218
QY      741  TGTGACTCCACAGACTCAAAAGGACTCCAGTCTTGAAGGCTGGACCTGGGGATGGGTT 800
Db      217  TGTGACTCCACAGACTCAAAAGGACTCCAGTCTTGAAGGCTGGACCTGGGGATGGGTT 158
QY      801  TCTCACACCCATATGTCTGTCCCTTGGATAGGCTGAGGCTGAAGCACAGGAGAAAT 860
Db      157  TCTCACACCCATATGTCTGTCCCTTGGATAGGCTGAGGCTGAAGCACAGGAGAAAT 98
QY      861  ATGTGCTTTCTTCTGCGCCCTACCTCTCTTCCATCTAGACTGTCTCTTGAGCCAGGCTCTG 920
Db      97  ATGTGCTTTCTTCTGCGCCCTACCTCTCTTCCATCTAGACTGTCTCTTGAGCCAGGCTCTG 38
QY      921  TAAACCTGCACATTTATATGTGTTTCACACATGTAAAGT 957
Db      37  TAAACCTGCACATTTATATGTGTTTCACACATGTAAAGT 1

RESULT 10
ABQ56941
ID      ABQ56941 standard; cDNA; 559 BP.
XX      AC      ABQ56941;
XX      DT      02-AUG-2002 (first entry)
XX      DE      Human colon cancer related nucleotide sequence SEQ ID NO:636.
XX      KW      Human; colon cancer; cancer; tissue profiling; forensic; mapping;
XX      KW      genetic analysis; diagnostic; antisense therapy; gene; ss.
XX      OS      Homo sapiens.
XX      PN      WO200229086-A2.
XX      PD      11-APR-2002.
XX      PF      02-OCT-2001; 2001WO-US030732.
XX      PR      02-OCT-2000; 2000US-0237271P.
XX      PA      (FARB ) BAYER CORP.
XX      PI      Burgess C, Astle JH, Carroll E, Catino TU, Dwivedi P, Molino GA;
XX      PI      Thiaglingam A, Lewis MB;
XX      DR      WPI; 2002-426115/45.
XX      PT      New isolated nucleic acid that is differentially expressed in cancer
XX      PT      tissues useful for determining the presence of colon cancer in a cell or
XX      PT      tissue type, and in antisense therapy.

```

PS	Claim 1; Fig 1; 796pp; English.	
XX	ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially expressed in cancer tissues. ABQ78993 to ABQ79004 represent proteins encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be used in antisense therapy. An antibody immunoreactive with a polypeptide encoded by (I) is useful for detecting cancer in a patient sample, and for detecting the presence or absence of a polynucleotide encoded by a nucleic acid which hybridizes to (I) in a cell. A probe/primer derived from (I) can be used for determining the presence of a nucleic acid which hybridizes to (I), and for determining the phenotype of cells in a sample of cells from a patient. (I) is useful for determining the presence of colon cancer in a cell or tissue type, for determining the presence or state of other type of cancer, in antisense therapy, to generate macroarrays on a solid surface, to identify a chromosome on which the corresponding gene resides, and in tissue profiling, forensics, genetic analysis, mapping and diagnostic applications. (I) can be used to raise antibodies, and to screen for peptide analogues and antagonists	
XX	Sequence 559 BP; 135 A; 140 C; 167 G; 117 T; 0 U; 0 Other;	
QY	Query Match 19.8%; Score 515.8; DB 6; Length 559;	
DB	Best Local Similarity 99.6%; Pred. No. 3.8e-96;	
QY	Matches 517; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
DB	109 CTGGGCTACTAGATGGCGATGATTTTCAGTGGCGCTGGCAGTATCGCTTCCACCCCTT 168	
QY	16 CTGGGCTACTAGATGGCGATGATTTTCAGTGGCGCTGGCAGTATCGCTTCCACCCCTT 75	
DB	169 CTTTAGTTTACACCAATGTGGACACTCGGCAGAGCAGCTGGCGCTGGTCTCGCT 228	
QY	76 CTTTAGTTTACACCAATGTGGACACTCGGCAGAGCAGCTGGCGCTGGTCTCGCT 135	
DB	229 GGTCTCTGCTCTTCTGCGGCTGCACAAACAGTCCAGCATGACGGTATGGAAGCTCAGGA 288	
QY	136 GGTCTCTGCTCTTCTGCGGCTGCACAAACAGTCCAGCATGACGGTATGGAAGCTCAGGA 195	
DB	289 GAGCCGCTCTTCAACAAAGTCAAGCTACAGCGAAAGCTTCTCTGTGGATCGATCCAGAT 348	
QY	196 GAGCCGCTCTTCAACAAAGTCAAGCTACAGCGAAAGCTTCTCTGTGGATCGATCCAGAT 255	
DB	349 TGTATTAGGAGTCTGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 408	
QY	256 TGTATTAGGAGTCTGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 315	
DB	409 CTTCTCTGATCTGTGGCGAGGCCACAGAAATGGGGGAACTCATCTATCAGTGGGTTTC 468	
QY	316 CTTCTCTGATCTGTGGCGAGGCCACAGAAATGGGGGAACTCATCTATCAGTGGGTTTC 375	
DB	469 CAGGAGTGCCCGAGAACTCCGCTCTTTTACCCTGTATGAACCTGACTTAATGGGGAAGACAC 528	
QY	376 CAGGAGTGCCCGAGAACTCCGCTCTTTTACCCTGTATGAACCTGACTTAATGGGGAAGACAC 435	
DB	529 AGAGGATGAGGAGTTTCAGGGCTGATGAGGCACTCTACTCGGGCTCTGAGGCCCT 588	
QY	436 AGAGGATGAGGAGTTTCAGGGCTGATGAGGCACTCTACTCGGGCTCTGAGGCCCT 495	
DB	589 ACAGCAGGAGCACAAGCGCGAGATCATCTGTGACCGA 627	
QY	496 ACAGCAGGAGCACAAGCGCGAGATCATCTGTGACCGA 534	
DB		
DE	RESULT 11	
XX	ADQ50360	
XX	ID ADQ50360 standard; DNA; 675 BP.	
XX	AC ADQ50360;	
XX	21-OCT-2004 (first entry)	
DE	Novel canine microarray-related DNA sequence SeqID1662.	
XX	canine microarray; drug screening; toxicity assay;	
QY		
DB		
QY	environmental pollutant; cellular response; gene expression profile; toxic response; liver necrosis; fatty liver disease; protein adduct formation; hepatitis; dog; ds.	
DB	Canis familiaris.	
QY	WO2004063324-A2.	
DB	29-JUL-2004.	
QY	05-MAY-2003; 2003WO-US013853.	
DB	03-MAY-2002; 2002US-0377240P.	
QY	(GENE-) GENE LOGIC INC.	
DB	(PFIZ) PFIZER PROD INC.	
QY	Diggans JC, Porter M, Wei T;	
DB	WPI; 2004-561890/54.	
QY	New isolated nucleic acid molecule, useful for drug screening and toxicity assays or for assessing the impact, including toxicity, of a compound, pharmaceutical agent or environmental pollutant on a cell or living organism.	
DB	Claim 1; SEQ ID NO 1662; 41pp; English.	
QY	This invention is related to a novel isolated canine nucleic acid sequences and the construction of canine microarrays containing a significant portion of the canine genome. The isolated canine nucleic acid sequences of the invention may be useful for drug screening and toxicity assays. The invention is therefore useful for assessing the impact, including toxicity, of a compound, pharmaceutical agent or environmental pollutant on a cell or living organism. The methods are useful for detecting genes that are up- or down-regulated in canines in a disease state. The sequences are useful as diagnostic agents or markers to detect a cellular response in a sample individually or as part of a gene expression profile. It is also useful as a target for agents that modulate gene expression or activity. The database is useful for producing electronic Northern that allow the user to determine the cell type or tissue in which a given gene is expressed and to allow determination of the abundance or expression level of a given gene in a particular tissue or cell. The methods are useful for determining the similarity of a toxic response to one or more individual compounds. The methods are useful for predicting at least one toxic response or the likelihood that a compound or test agent will induce various specific pathologies such as those of the liver (liver necrosis, fatty liver disease, protein adduct formation or hepatitis), those of the kidney, heart, brain or testes, or other pathologies associated with at least one of the toxins. The methods are also useful for predicting or elucidating the potential cellular pathways influenced, induced or modulated by the compound or test agent due to the similarity of the expression profile compared to the profile induced by a known toxin. The present sequence is that of a canine DNA sequence which was claimed for use during the production of a canine microarray of the invention.	
DB	Sequence 675 BP; 159 A; 189 C; 170 G; 155 T; 0 U; 2 Other;	
QY	Query Match 18.8%; Score 489.2; DB 13; Length 675;	
DB	Best Local Similarity 86.2%; Pred. No. 1.2e-90;	
QY	Matches 595; Conservative 0; Mismatches 70; Indels 25; Gaps 4;	
DB	433 AGAAGAATGGGGAACTCATCTATCAGTGGGTTTCCAGAGTGGCCAGAACAACTCCGT 492	
QY	1 AGAAGAATGGGGAACTCATCTATCAGTGGGTTTCCAGAGTGGCCAGAACAACTCCGT 60	
DB	493 CTTTACCCCTGTATGAACCTGACTTAATGGGGAAGACACAGAGGATGAGGATTCACGGCT 552	
QY	61 GTTCACCCCTGTATGAACCTGACTTAATGGGGAAGACACAGAGGATGAGGATTCACGGCT 120	
DB	553 GGATGAGCCACTTACTCGGGCTCTGAGGCCCTTACAGCAGGAGCAGACAGGCCAGAT 612	
QY		

Db 121 GGATGAAGCAACCTTACTTGGGCTCTGAGGCTCTGAGAT 180
 Qy 613 CATCACTGTGACGATGCGGAGGCTCAAGTCTTCTAGCAGGACCTGTCTCCCTTTA 672
 Db 181 CATCACTGTGACGATGCGGAGGCTCAAGTCTTCTAGCAGGACCTGTCTCTC---TTA 237
 Qy 673 CTTCTTACTCTCCACTCTTTCAGAGGCTTCAAAAGGAGACAGACCCAGTGTCTCCCAAG 732
 Db 238 CTTCTTACTCTCCACTCTTTCAGAGGCTTCAAAAGGAGACAGACCCAGTGTCTCCAG 296
 Qy 733 ACTGGATCTGACTCCACAGACTCAAAAGGACTCCAGTCTCTGAAGCTGGGACTGGG 792
 Db 297 ACTGGATCTGACTCCACAGACTCAAAAGGCTCCAGT-----CCAGGG 342
 Qy 793 GATGGGTTTCTCACACCCCATATGCTGTCTCTTGGATAGGCTGAGGCTGAAGCACCAGG 852
 Db 343 ATGGGTTTCCCACTACCTGTATGTCTGTCTTGGGAAAGTGAGGCTGAGGTACCCAGG 402
 Qy 853 GAGAAATATGTGCTTCTTCTGCGCTTACCTCTTCCATCTCTAGACTGTCTCTTGAGCC 912
 Db 403 GAGAAAGATGTCTTCTTCTGCTGCTCTCTCTACCATCTTACATCTGTCTCTGAGCC 462
 Qy 913 AGGGTCTGTAACTGTGACATTTATATGTGTTCACATGTGAAGTACATACATGCG 972
 Db 463 AGGGTCTGTAACTTTAACTTCAATGTGTTCATACATGTTTCAATAGTAAATACATAC 515
 Qy 973 CTGCGACGACATGCTTCT 1032
 Db 516 GCGCTGTGACACAGCTTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 575
 Qy 1033 TCTCAGGCTGTGCTGGATCTCTCTCTAGGGATGGGGAAGCCCTGGCTGACGGCAGGCT 1092
 Db 576 TCTCAGGCTGTGCTGGATCTCTCTCTAGGGATGGGGAAGCCCTGGCTGACGGCAGGCT 635
 Qy 1093 TCCAGGCAATATGAAGTAGGAGGCCACG 1122
 Db 636 TCCAGGCAATATGAAGTAGGAGGCCACG 665

RESULT 12
 ABQ56838/c
 ID ABQ56838 standard; cDNA; 507 BP.
 AC ABQ56838;
 XX
 DT 02-AUG-2002 (first entry)
 DE Human colon cancer related nucleotide sequence SEQ ID NO:533.
 KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;
 KW genetic analysis; diagnostic; antisense therapy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200229086-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 02-OCT-2001; 2001WO-US030732.
 XX
 PR 02-OCT-2000; 2000US-0237271P.
 XX
 PA (FARB) BAYER CORP.
 XX
 PI Burgess C, Aetle JH, Carroll E, Catino TU, Dwivedi P, Molino GA;
 PI Thiallingam A, Lewis ME;
 XX
 DR WPI; 2002-426115/45.
 XX
 PT New isolated nucleic acid that is differentially expressed in cancer
 PT tissues useful for determining the presence of colon cancer in a cell or
 PT tissue type, and in antisense therapy.
 XX

PS Claim 1; Fig 1; 796pp; English.
 XX
 CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
 CC expressed in cancer tissues. ABQ78993 to ABQ79004 represent proteins
 CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
 CC used in antisense therapy. An antibody immunoreactive with a polypeptide
 CC encoded by (I) is useful for detecting cancer in a patient sample, and
 CC for detecting the presence or absence of a polynucleotide encoded by a
 CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
 CC from (I) can be used for determining the presence of a nucleic acid which
 CC hybridises to (I), and for determining the phenotype of cells in a sample
 CC of cells from a patient. (I) is useful for determining the presence of
 CC colon cancer in a cell or tissue type, for determining the presence or
 CC state of other type of cancer, in antisense therapy, to generate the
 CC macroarrays on a solid surface, to identify a chromosome on which the
 CC corresponding gene resides, and in tissue profiling, forensics, genetic
 CC analysis, mapping and diagnostic applications. (I) can be used to raise
 CC antibodies, and to screen for peptide analogues and antagonists
 XX
 SQ Sequence 507 BP; 123 A; 125 C; 133 G; 120 T; 0 U; 6 Other;
 Query Match 17.8%; Score 464.4; DB 6; Length 507;
 Best Local Similarity 97.0%; Pred. No. 1.3e-85;
 Matches 491; Conservative 0; Mismatches 12; Indels 3; Gaps 2;
 Qy 455 TATCAGTGGGTTTCCAGGA--GTGGCCAGAACTCCGTCTTTACCTGTATGAAC-TG 511
 Db 506 TATCAGTNGTTCAGGAGTGGCCAGAACACTNGTCTTTACCTGTATGAAC-TTG 447
 Qy 512 ACTAATGGGGAAGACACAGAGGATGAGAGTTCCACGGGCTGGATGAAGCCACTTACTG 571
 Db 446 ACTAATGGGGAAGACACAGAGGATGAGAGTTCCACGGGCTGGATGAAGCCACTTACTG 387
 Qy 572 CGGGCTCTGAGGCTTACAGAGGACACAGGCGGAGATCATCTGTCAGCGATGCG 631
 Db 386 CGGGCTCTGAGGCTTANAGCGAGGAGCAAGGCGGAGATCATCANTGTGAGCGATG 327
 Qy 632 CGAGGCTCAAGTCTTCTAGCAGGACCTGTCTCTCTTACTTCTTACCTCCACCTTT 691
 Db 326 CGAGGCTCAAGTCTTCTAGCAGGACCTGTCTCTCTTACTTCTTACCTCCACCTTT 267
 Qy 692 CCAGGCTTTCAAAGGAGACAGACCCAGTGTCCCCCAAAGACTGGATCTGTGATCCAC 751
 Db 266 CCAGGCTTTCAAAGGAGACAGACCCAGTGTCCCCCAAAGACTGGATCTGTGATCCAC 207
 Qy 752 CAGACTCAAAGGACTCCAGTCTTCAAGGCTGGGAGCTGGGATGGTCTTCACACCCC 811
 Db 206 CAGACTCAAAGGACTCCAGTCTTCAAGGCTGGGAGCTGGGATGGTCTTCACACCCC 147
 Qy 812 ATATGCTGTCTTGGATAGGCTGAGGCTGAAGCACCAGGAGAGAAATATGTGCTTCTT 871
 Db 146 ATATGCTGTCTTGGATAGGCTGAGGCTGAAGCACCAGGAGAGAAATATGTGCTTCTT 87
 Qy 872 CTGCGCTTACCTCTCTTCCATCTCTAGACTGTCTCTTGAAGCCAGGCTGTGAACTGACA 931
 Db 86 CTGCGCTTACCTCTCTTCCATCTCTAGACTGTCTCTTGAAGCCAGGCTGTGAACTGACA 27
 Qy 932 CTTTATATGTGTTTTCACATGTAAGT 957
 Db 26 CTTTATATGTGTTTTCACATGTAAGT 1
 RESULT 13
 ABQ56915
 ID ABQ56915 standard; cDNA; 609 BP.
 XX
 AC ABQ56915;
 XX
 DT 02-AUG-2002 (first entry)
 DE Human colon cancer related nucleotide sequence SEQ ID NO:610.
 XX
 KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;

genetic analysis; diagnostic; antisense therapy; gene; ss.
Homo sapiens.
WO200229086-A2.
11-APR-2002.
02-OCT-2001; 2001WO-US030732.
02-OCT-2000; 2000US-0237271P.
(PARB) BAYER CORP.
Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA, Thiagalingam A, Lewis ME;
WPI; 2002-426115/45.
New isolated nucleic acid that is differentially expressed in cancer tissues useful for determining the presence of colon cancer in a cell or tissue type, and in antisense therapy.
Claim 1; Fig 1; 796pp; English.
ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially expressed in cancer tissues. ABB78993 to ABB79004 represent proteins encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be used in antisense therapy. An antibody immunoreactive with a polypeptide encoded by (I) is useful for detecting cancer in a patient sample, and for detecting the presence or absence of a polynucleotide encoded by a nucleic acid which hybridizes to (I) in a cell. A probe/primer derived from (I) can be used for determining the presence of a nucleic acid which hybridizes to (I), and for determining the phenotype of cells in a sample of cells from a patient. (I) is useful for determining the presence of colon cancer in a cell or tissue type, for determining the presence or state of other type of cancer, in antisense therapy, to generate macroarrays on a solid surface, to identify a chromosome on which the corresponding gene resides, and in tissue profiling, forensics, genetic analysis, mapping and diagnostic applications. (I) can be used to raise antibodies, and to screen for peptide analogues and antagonists
Sequence 609 BP; 156 A; 149 C; 160 G; 135 T; 0 U; 9 Other;
Query Match 15.2%; Score 395; DB 6; Length 609;
Best Local Similarity 96.5%; Pred. No. 2.6e-71;
Matches 447; Conservative 0; Mismatches 10; Indels 6; Gaps 4;
112 GGCTACTAGTGGCGATGAGTTTCGAGTGGCGGTGGCGATATCGCTTCCACCTTCTT 171
|||
5 GGGGACTACGATGGTGTGATGTTTCGAGTGGCGGTGGCGATATCGCTTCCACCTTCTT 64
172 TACGTTACACCGATGTGGACACTCGGCAGAGCAGCTGGCGGCTGGTGGCTCGCTGGT 231
|||
65 TACGTTACACCGATGTGGACACTCGGCAGAGCAGCTGGCGGCTGGTGGCTCGCTGGT 124
232 CCTGTCTTCTCCGCTGTCACAAACAGTCCAGCATGACGGTGTGGAGTCCAGAGAG 291
135 CCTGTCTTCTCCGCTGTCACAAACAGTCCAGCATGACGGTGTGGAGTCCAGAGAG 184
292 CCCGCTCTTCAACACGTCACAGTACAGCGAAGCTTCTGTGGAGTGCATCCAGATTGT 351
185 CCCGCTCTTCAACACGTCACAGTACAGCGAAGCTTCTGTGGAGTGCATCCAGATTGT 244
352 ATTAGGAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 411
245 ATTAGGAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 304
412 CCTGTATCATGTGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 471
305 CCTGTATCATGTGGCGAG-GCCAGAGAGATGGGGGAGAACTCATTCATCATGTGG--TTTCAG 361
472 GAGTGCCGAGAGAACTCCGCTTTTACCTGTATGAACTGACTATATGGGGAGACACAGA 531

Db 362 GAGTGGC--AGAACACACTCGCTCTTTTACCTGTATGAACCTACTTAATGGGAGACACAGA 419
QY 532 GGATGAGGAGTTCCACGGCTGGATGAAGCCACTCTTACTGCGG 574
Db 420 GGATGAGGAGTTCCACGGCTGGATGA--CCACTTACTGCGG 461
RESULT 14
ABX17514
ID ABX17514 standard; cDNA; 545 BP.
XX AC ABX17514;
XX 05-FEB-2003 (first entry)
XX Human cDNA encoding protein NOV5.
XX Human; ss; gene; NOV5; adrenoleukodystrophy; haemophilia; stoke; VHL;
KW congenital adrenal hyperplasia; haemophilia; hypercoagulation;
KW idiopathic thrombocytopenic purpura; autoimmune disease; allergy;
KW immunodeficiencies; transplantation; Von Hippel-Lindau syndrome;
KW Alzheimer's disease; tuberculous sclerosis; Parkinson's disease; epilepsy;
KW Huntington's disease; cerebral palsy; Lesch-Nyhan syndrome; pain;
KW multiple sclerosis; ataxia-telangiectasia; leukodystrophy; anxiety;
KW behavioural disorder; addiction; neuroprotection; diabetes; ARDS;
KW renal artery stenosis; interstitial nephritis; glomerulonephritis;
KW polycystic kidney disease; systemic lupus erythematosus; IGA;
KW renal tubular acidosis; immunoglobulin A nephropathy; hypercalcaemia;
KW cirrhosis; transplantation; asthma; emphysema; scleroderma; GWD;
KW adult respiratory distress syndrome; graft versus host disease;
KW lymphedema; fertility; pancreatitis; obesity; haemophilia; ulcer;
KW anaemia; cancer; trauma; regeneration; infection.
XX Homo sapiens.
OS WO2002B1629-A2.
XX 17-OCT-2002.
XX 03-APR-2002; 2002WO-US010522.
XX 03-APR-2001; 2001US-0281086P.
PR 03-APR-2001; 2001US-0281136P.
PR 05-APR-2001; 2001US-0281863P.
PR 05-APR-2001; 2001US-0281906P.
PR 06-APR-2001; 2001US-0282020P.
PR 10-APR-2001; 2001US-0282934P.
PR 12-APR-2001; 2001US-0283512P.
PR 19-APR-2001; 2001US-0285325P.
PR 23-APR-2001; 2001US-0285890P.
PR 24-APR-2001; 2001US-0286068P.
PR 25-APR-2001; 2001US-0286292P.
PR 27-APR-2001; 2001US-0287213P.
PR 02-MAY-2001; 2001US-0288257P.
PR 12-MAY-2001; 2001US-0291134P.
PR 31-MAY-2001; 2001US-0291725P.
PR 08-JUN-2001; 2001US-0296965P.
PR 18-JUN-2001; 2001US-0299128P.
PR 12-JUL-2001; 2001US-0305063P.
PR 14-NOV-2001; 2001US-0332780P.
PR 04-JAN-2002; 2002US-0345221P.
PR 02-APR-2002; 2002US-00345221.
XX (CURA-) CURAGEN CORP.
XX Spytek KA, Li L, Edinger SR, Ellerman K, Stone DJ, Malyankar UM;
PI Shinkets RA, Guo X, Anderson DW, Patturajan M, Berghs C, Gerlach V;
PI Taupier RJ, Pena CE, Padigaru M, Liu Y, Burgess CE, Miller CE;
PI Gusev VV, Kekuda R, Gorman L, Zernhusen BD, Baumgartner JC;
PI Tchernev VT, Vernet CAM, Smithson G, Heyes MP, Shenoy SG, Liu X;
PI Gangolli EA;

XX WPI: 2003-046863/04.
 DR P-PSDB; ABU11036.
 XX
 PT New polypeptides, designated NOVX polypeptides, useful for treating
 PT hemophilia, idiopathic thrombocytopenic purpura, autoimmune disease,
 PT allergies, transplantation, Alzheimer's disease and stroke.
 XX
 PS Claim 19; Page 96-97; 320pp; English.
 XX
 CC The invention relates to an isolated NOVX polypeptide selected from NOV1-
 CC NOV27 polypeptides, a mature form of NOVX, a variant of NOVX or a
 CC fragment of NOVX. Also included are determining the presence or amount of
 CC NOVX in a sample (by using an antibody that immunospecifically bind to
 CC the polypeptide), determining the presence of or predisposition to
 CC disease associated with altered levels of NOVX in a first mammalian
 CC subject, identifying a potential therapeutic agent for use in the
 CC treatment of pathology related to aberrant expression of physiological
 CC interactions of NOVX, screening for a modulator of activity or of latency
 CC or predisposition to a pathology associated with NOVX, the nucleic acid
 CC encoding NOVX, vectors and host cells. NOVX is useful for identifying an
 CC agent (a cellular receptor or downstream effector) that binds to NOVX.
 CC NOVX and NOVX nucleic acids are useful for treating or preventing NOVX-
 CC associated disorders in humans, and in the manufacture of a medicament
 CC for treating a NOVX related disease human disease e.g.
 CC adrenoleukodystrophy, congenital adrenal hyperplasia, haemophilia,
 CC hypercoagulation, idiopathic thrombocytopenic purpura, autoimmune
 CC disease, allergies, immunodeficiencies, transplantation, Von Hippel-
 CC Lindau (VHL) syndrome, Alzheimer's disease, stroke, tuberosus sclerosis,
 CC Parkinson's disease, Huntington's disease, cerebellar palsy, epilepsy,
 CC Lesch-Nyhan syndrome, multiple sclerosis, ataxia-telangiectasia,
 CC leukodystrophies, behavioural disorders, addiction, anxiety, pain,
 CC neuroprotection, diabetes, renal artery stenosis, interstitial nephritis,
 CC glomerulonephritis, polycystic kidney disease, systemic lupus
 CC erythematosus, renal tubular acidosis, immunoglobulin (Ig) A nephropathy,
 CC hypercalcaemia, cirrhosis, transplantation, asthma, emphysema,
 CC scleroderma, adult respiratory distress syndrome (ARDS), graft versus
 CC host disease (GVHD), lymphedema, fertility, pancreatitis, obesity,
 CC haemophilia, ulcers, anaemia, cancer, trauma, regeneration, and viral,
 CC bacterial or parasitic infections. The present sequence encodes a NOVX
 CC protein
 XX
 SQ Sequence 545 BP; 135 A; 117 C; 159 G; 134 T; 0 U; 0 Other;
 Query Match 13.6%; Score 353.6; DB 8; Length 545;
 Best Local Similarity 84.6%; Pred. No. 8.2e-63;
 Matches 435; Conservative 0; Mismatches 69; Indels 10; Gaps 3;
 QY 109 CTGGGCTACTACGATGGCGATGAGTTTCGAGTGGCGCGTTCGCCAGTATCGTTCGCCACCCTT 168
 DB 8 CTGGGCTATTATGATGGTGACGAATTCAGTGGTGTGGGAGTATCGTTCGCCGCC--- 64
 QY 169 CTTTACGTTTACACCGAATGTGACACTGGGAGAGAGAGTGGCGCGCTGGTGTCTGCT 228
 DB 65 CTTTACATACAGCTGAGCTGGCCACTGGGAGAGAGAGTGGCCACTGGTGTGTTGTT 124
 QY 229 GGTCTGTCTTCTCGCGCTGCACAAACAGTCCAGCATGACGGTGTGAGTCCAGGACGAG 288
 DB 125 GGTCTGTCTTCTCGCGCTGCACAAACAGTCCAGCATGATGGTGTGAGTGTCTCAGGA 184
 QY 289 GAGCCCGCTCTTCAACAGCTCAAGTACAGGAGAAAGCTTCTGTGGAGTGCATCCAGAT 348
 DB 195 GATCTGTCTTTCAGCAACATCAAGCT---GTGGAGAGTCTTCTGTGGATCATCCAGT 241
 QY 349 TGTATTAGAGGAACGTGAGGAGAAAGAGGAACTCTCGAGTGGTGTGGATTAAGAGCAAGTCCAG 408
 DB 242 TGTATTAGAGGAACGTGAGGAGAAAGAGGAACTCTCGAGTGGTGTGGATTAAGAGCAAGTCTAG 301
 QY 409 CTTCTGTATGATGTGGCGAGCCAGAGAGATGGGGGAACTCATCTATCAGTGGGTTTC 468
 DB 302 TTTCTTAATCATGTGGCGAGGCCAGAGAAATGGGGGAAACTCATCTATCAGTGGGTTTC 361
 QY 469 CAGGAGTGGCCAGAACAACTCGCTCTTTACCTGTATGAACCTGACTAATGGGGAAGACAC 528

DB 362 CAGGAGTGGCCAGAACAACTCGTACTTAGCTGTATGAGCTGACCAATGGGGAAGACAT 421
 QY 529 AGAGATGAGGAGTTCACGGGCTGGATGAAGACCACTCTACTGCGGGCTCTGCAGGCCCT 588
 DB 422 AGAGAAATGAGGTGTTCACGGGCTAAAGAGGCG---CTTCTGTGGGCTCTGCAGGCCCT 477
 QY 589 ACAGCAGGACGACACGAGCGGAGATCATCACTGTC 622
 DB 478 TCAGTAGGAACATAGGCTGAGATCATCACCATC 511
 RESULT 15
 ID ABQ56970 standard; cDNA; 349 BP.
 AC ABQ56970;
 XX ABQ56970;
 DT 02-AUG-2002 (first entry)
 DE Human colon cancer related nucleotide sequence SEQ ID NO:665.
 XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;
 KW genetic analysis; diagnostic; antisense therapy; gene; ss.
 XX Homo sapiens.
 OS WO200229086-A2.
 PN 11-APR-2002.
 PD 02-OCT-2001; 2001WO-US030732.
 PF 02-OCT-2000; 2000US-0237271P.
 PR (FARB) BAYER CORP.
 PA Burgess C, Aetle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
 PI Thiaglingam A, Lewis ME;
 XX WPI; 2002-426115/45.
 XX New isolated nucleic acid that is differentially expressed in cancer
 PT tissues useful for determining the presence of colon cancer in a cell or
 PT tissue type, and in antisense therapy.
 XX Claim 1; Fig 1; 796pp; English.
 CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
 CC expressed in cancer tissues. ABQ78993 to ABQ79004 represent proteins
 CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
 CC used in antisense therapy. An antibody immunoreactive with a polypeptide
 CC encoded by (I) is useful for detecting cancer in a patient sample, and
 CC for detecting the presence or absence of a polynucleotide encoded by a
 CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
 CC from (I) can be used for determining the presence of a nucleic acid which
 CC hybridises to (I), and for determining the phenotype of cells in a sample
 CC of cells from a patient. (I) is useful for determining the presence of
 CC colon cancer in a cell or tissue type, for determining the presence or
 CC state of other type of cancer, in antisense therapy, to generate
 CC macroarrays on a solid surface, to identify a chromosome on which the
 CC corresponding gene resides, and in tissue profiling, forensics, genetic
 CC analysis, mapping and diagnostic applications. (I) can be used to raise
 CC antibodies, and to screen for peptide analogues and antagonists
 XX
 SQ Sequence 349 BP; 75 A; 102 C; 82 G; 86 T; 0 U; 4 Other;
 Query Match 12.2%; Score 317.4; DB 6; Length 349;
 Best Local Similarity 98.5%; Pred. No. 2e-55;
 Matches 318; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 637 CGTCAAGTTCTTCTAGCAGGACCTGTCTCCCTTTACTTCTTACCTCCACCTTCCAGG 696

Db	1	CGTCAAGTTCTTCTAGCAGGACCTGTCTCCCTTTACTTCTTACCTCCACCTTTCCAGG	60
Qy	697	GCTTTCAAAAGGAGACAGACCCAGTGTCCCCAAAGACTGGATCTGTGACTCCACCAGAC	756
Db	61	GCTTTCAAAAGGAGACAGACCCAGTGTCCCCAAAGACTGGATCTGTGACTCCACCAGAC	120
Qy	757	TCAAAAGGACTCCAGTCTCTGAAGGCTGGGACCTGGGGATGGGTTCTCACACCCCATATG	816
Db	121	TCAAAAGGACTCCAGTCTCTGAAGGCTGGGACCTGGGGATGGGTTCTCACACCCCATATG	180
Qy	817	TCTGTCCCTTGGATAGGGTGAGGCTGAAGCACCCAGGGAGAAAATATGTCTTCTCTCGC	876
Db	181	TCTGTCCCTTGGATAGGGTGAGGCTGAAGCACCCAGGGAGAAAATATGTCTTCTCTCGC	240
Qy	877	CCTACCTCCTTTCCCATCCTAGACTGTCTTGGAGCCAGGGTCTGTAAACCTGACACTTTA	936
Db	241	CCTACCTCCTTTCCCATCCTAGACTGTCTTGGAGCCAGGGTCTGTAAACCTGACACTTTA	300
Qy	937	TATGTGTTCCACATGTAGTAC	959
Db	301	TATGTGTTCCACATGTAGTAC	323

Search completed: June 28, 2005, 19:33:46
 Job time : 918.543 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 18:22:52 ; Search time 295.707 Seconds
(without alignments)
14414.621 Million cell updates/sec

Title: US-08-731-499-4
Perfect score: 2605
Sequence: 1 CAAGCTCGAATAACCTC.....CTGGCGTTACCACTTAAT 2605

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B COMB.seq: *
5: /cgn2_6/prodata/1/ina/PTUS COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2605	100.0	2605	2	US-08-680-395-4
2	2605	100.0	2605	4	US-08-892-693-4
3	1076.2	41.3	1243	4	US-09-620-312D-930
4	476	18.3	495	4	US-09-621-976-308
5	453	17.4	487	4	US-09-621-976-307
6	444	17.0	494	4	US-09-621-976-1902
7	389.6	15.0	417	4	US-09-621-976-15507
8	386.4	14.8	400	4	US-09-621-976-15511
9	362	13.9	401	4	US-09-621-976-15506
10	328.2	12.6	383	4	US-09-621-976-15508
11	309	11.9	360	4	US-09-621-976-15509
12	269	10.3	269	4	US-09-702-705-1380
13	269	10.3	269	4	US-09-736-457-1380
14	269	10.3	269	4	US-09-614-124B-1380
15	269	10.3	269	4	US-09-671-325-1380
16	269	10.3	269	4	US-09-658-824-1380
17	237	8.7	241	4	US-09-471-276-611
18	209	8.0	261	4	US-09-621-976-15510
19	190	7.3	195	4	US-09-513-999C-433
20	156.4	6.0	1763	4	US-09-244-805-57
21	156.4	6.0	1824	4	US-09-244-805-5
22	145	5.6	5592	3	US-09-495-797-37
23	145	5.6	12479	4	US-09-318-138-13
24	145	5.6	12494	3	US-08-935-312-13
25	145	5.6	12494	3	US-08-848-760B-33
26	139.4	5.4	3219	4	US-09-509-800-3
27	138	5.3	4525	2	US-08-613-861-2

28	137.4	5.3	3385	4	US-09-509-800-1	Sequence 1, Appli
29	136	5.2	3699	3	US-08-646-538-6	Sequence 6, Appli
30	136	5.2	3699	3	US-09-503-222-6	Sequence 6, Appli
31	136	5.2	4965	3	US-08-675-566-22	Sequence 22, Appli
32	135	5.2	9299	3	US-09-097-319A-15	Sequence 15, Appli
33	135	5.2	9299	4	US-09-643-971-15	Sequence 15, Appli
34	134	5.1	6596	3	US-09-575-602-11	Sequence 11, Appli
35	134	5.1	6596	4	US-09-032-086-11	Sequence 12, Appli
36	133.8	5.1	3988	4	US-09-358-856C-12	Sequence 10881, A
37	133	5.1	902	4	US-09-270-767-10881	Sequence 1, A
38	133	5.1	6994	3	US-08-675-566-1	Sequence 1, Appli
39	133	5.1	7001	3	US-08-675-566-3	Sequence 3, Appli
40	133	5.1	9335	3	US-09-097-319A-19	Sequence 19, Appli
41	133	5.1	9335	4	US-09-643-971-19	Sequence 19, Appli
42	133	5.1	9408	3	US-09-097-319A-16	Sequence 16, Appli
43	133	5.1	9408	4	US-09-643-971-16	Sequence 16, Appli
44	130	5.0	4883	3	US-09-608-730B-21	Sequence 21, Appli
45	127.6	4.9	713	3	US-08-998-416-135	Sequence 135, App

ALIGNMENTS

RESULT 1
US-08-680-395-4
; Sequence 4, Application US/08680395
; Patent No. 5892010
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W.
; APPLICANT: Collins, Colin
; APPLICANT: Hwang, Soo-in
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowbel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: Genes from the 20q13 Amplicon and Their
; TITLE OF INVENTION: Uses
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/680,395
; FILING DATE: 15-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-0689000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2605 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..2605
; OTHER INFORMATION: /note= "cdna clone cc43 of 4 kb
; OTHER INFORMATION: transcript"
US-08-680-395-4

Query Match									
Best Local Similarity 100.0%; Score 2605; DB 2; Length 2605;									
Matches 2605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	CAAGCTCGAAATTAACCCCTCACTAAAGGAAACAAAGCTGAGCTCCACCGCGGTGGCGG	60						
DB	1	CAAGCTCGAAATTAACCCCTCACTAAAGGAAACAAAGCTGAGCTCCACCGCGGTGGCGG	60						
QY	61	CGGCTCTAGAACTAGTGGATCCCGCGGCTGAGGAAATTCGGCAGAGCTGGGCTACTAC	120						
DB	61	CGGCTCTAGAACTAGTGGATCCCGCGGCTGAGGAAATTCGGCAGAGCTGGGCTACTAC	120						
QY	121	GATGCGGATGAGTTCGAGTGGCGCTGGCAGTATCGCTTCCACCCCTCTTTACGTTACA	180						
DB	121	GATGCGGATGAGTTCGAGTGGCGCTGGCAGTATCGCTTCCACCCCTCTTTACGTTACA	180						
QY	181	ACCGAATGTGGACACTCGGACAGAGCTGGCGCCCTGGTGTCTCGCTGCTCTCTCT	240						
DB	181	ACCGAATGTGGACACTCGGACAGAGCTGGCGCCCTGGTGTCTCGCTGCTCTCTCT	240						
QY	241	CTGCGGCTGCAAAACAGTTCAGCATGACGCTGATGAGAGCTCAGGAGAGCCGCTCTT	300						
DB	241	CTGCGGCTGCAAAACAGTTCAGCATGACGCTGATGAGAGCTCAGGAGAGCCGCTCTT	300						
QY	301	CAACAACGTCAAGCTACAGCGAAAGCTTCTGTGAGTCCGATCCAGATTTGTTATAGAGA	360						
DB	301	CAACAACGTCAAGCTACAGCGAAAGCTTCTGTGAGTCCGATCCAGATTTGTTATAGAGA	360						
QY	361	ACTGAGGAAGAAAGGAACTCGAGTGGTTGATGAGAGCAAGTCCAGCTTCTCTGATCAT	420						
DB	361	ACTGAGGAAGAAAGGAACTCGAGTGGTTGATGAGAGCAAGTCCAGCTTCTCTGATCAT	420						
QY	421	GTGGCGGAGGCGAGAAATGGGGGAACTCATCTATACATGGGGTTCCAGAGTGGGCA	480						
DB	421	GTGGCGGAGGCGAGAAATGGGGGAACTCATCTATACATGGGGTTCCAGAGTGGGCA	480						
QY	481	GAACAACCTCCGTCTTACCTCTATGAACTGACTAATGGGGAAGACACAGAGATGAGGA	540						
DB	481	GAACAACCTCCGTCTTACCTCTATGAACTGACTAATGGGGAAGACACAGAGATGAGGA	540						
QY	541	GTTCACCGGCTGGATGAGCCACTCTACTGGGGCTCTGACGGCCCTACAGCAGGAGCA	600						
DB	541	GTTCACCGGCTGGATGAGCCACTCTACTGGGGCTCTGACGGCCCTACAGCAGGAGCA	600						
QY	601	CAAGCCGAGATCATCTGTGAGGATGAGCCACTCTACTGGGGCTCTGACGGCCCTACAGCAGGAGCA	660						
DB	601	CAAGCCGAGATCATCTGTGAGGATGAGCCACTCTACTGGGGCTCTGACGGCCCTACAGCAGGAGCA	660						
QY	661	TGTCTCCCTTTACTTTTACCTCCACCTTTTCCAGGGCTTTCAAAGGAGACAGACCCAG	720						
DB	661	TGTCTCCCTTTACTTTTACCTCCACCTTTTCCAGGGCTTTCAAAGGAGACAGACCCAG	720						
QY	721	TGTCCCCCAAGACTGGATCTGTGATCCACAGACTCAAAAGGACTCCAGTCTCTGAAGG	780						
DB	721	TGTCCCCCAAGACTGGATCTGTGATCCACAGACTCAAAAGGACTCCAGTCTCTGAAGG	780						
QY	781	CTGGACCTGGGGATGGGTTCTCACAACCCATATGTCTGCTCCCTTGGATAGGGTGGGC	840						
DB	781	CTGGACCTGGGGATGGGTTCTCACAACCCATATGTCTGCTCCCTTGGATAGGGTGGGC	840						
QY	841	TGAAGCACAGGAGAAATATGTCTTCTCTCCGCTTACCTCTCTTCCCATCTCTAGAC	900						
DB	841	TGAAGCACAGGAGAAATATGTCTTCTCTCCGCTTACCTCTCTTCCCATCTCTAGAC	900						
QY	901	TGTCTCTGAGCCAGGGTCTGTAACCTGACACTTTATATGTTTACACATGTAAGTACA	960						
DB	901	TGTCTCTGAGCCAGGGTCTGTAACCTGACACTTTATATGTTTACACATGTAAGTACA	960						
QY	961	TACACACATGCGGCTGACGACATGTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1020						
DB	961	TACACACATGCGGCTGACGACATGTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1020						

QY	1021	TGTTGCCCTCCCTTCTCAGCTGGTCTCGATCTCTCTAGGGGATGGGGAAGCCCTGCG	1080
DB	1021	TGTTGCCCTCCCTTCTCAGCTGGTCTCGATCTCTCTAGGGGATGGGGAAGCCCTGCG	1080
QY	1081	TGACGAGAGCCCTTCCAGGCAATATCAAGATAGGAGGCCCAACGGGCTGGCAGTGAAGT	1140
DB	1081	TGACGAGAGCCCTTCCAGGCAATATCAAGATAGGAGGCCCAACGGGCTGGCAGTGAAGT	1140
QY	1141	GTGGCCCAACACCGGATTTATGATATAAATCTCAATCTCCCAAAAAAAGAAAAA	1200
DB	1141	GTGGCCCAACACCGGATTTATGATATAAATCTCAATCTCCCAAAAAAAGAAAAA	1200
QY	1201	CTGAGACTAGTCTCTCTCTCGAGAACTAGTCTCGAGTCTTTTCTTTTCTTTT	1260
DB	1201	CTGAGACTAGTCTCTCTCTCGAGAACTAGTCTCGAGTCTTTTCTTTTCTTTT	1260
QY	1261	TTTTTTTTTTTTTTTTTTTTTAAAGATTTATTTATTTTCTCTCTCTCTCTCTCT	1320
DB	1261	TTTTTTTTTTTTTTTTTTTTTAAAGATTTATTTATTTTCTCTCTCTCTCTCTCT	1320
QY	1321	ACTTTTCTCTACTTAATACTACTTTCCAGTCTCAGAGCCCAAGGAAAAAAGAC	1380
DB	1321	ACTTTTCTCTACTTAATACTACTTTCCAGTCTCAGAGCCCAAGGAAAAAAGAC	1380
QY	1381	CATGAATCTTCTCTCCAGATTTAAAGTACACACTTTGGAAAAAGATTTGGAAAACTTT	1440
DB	1381	CATGAATCTTCTCTCCAGATTTAAAGTACACACTTTGGAAAAAGATTTGGAAAACTTT	1440
QY	1441	CTGAAAAAGTTGACTGAAACTCCAAACCAACATGCCATATTTGTTGATGTTCTCATGAA	1500
DB	1441	CTGAAAAAGTTGACTGAAACTCCAAACCAACATGCCATATTTGTTGATGTTCTCATGAA	1500
QY	1501	AATTTGTTAAACCTGTTCTAGATAAAGACAGTCTCAAGTTTGTACAGCCTACACAT	1560
DB	1501	AATTTGTTAAACCTGTTCTAGATAAAGACAGTCTCAAGTTTGTACAGCCTACACAT	1560
QY	1561	AGTACAAGGGTCCCTATGATGATTTCTCTGAGGACGAAATAAATGTAATTTTTCAGTT	1620
DB	1561	AGTACAAGGGTCCCTATGATGATTTCTCTGAGGACGAAATAAATGTAATTTTTCAGTT	1620
QY	1621	TCTGTTTTATAACTCTCTCGATCTCAGAGTTGACTGATTAACAACCTACTCATGCAACA	1680
DB	1621	TCTGTTTTATAACTCTCTCGATCTCAGAGTTGACTGATTAACAACCTACTCATGCAACA	1680
QY	1681	GAGATAAGGCACTCATATTTTATAAATATATGACCAAACTATTTTGGAAAACTTAT	1740
DB	1681	GAGATAAGGCACTCATATTTTATAAATATATGACCAAACTATTTTGGAAAACTTAT	1740
QY	1741	CTATTGGAGACACAAATATGCTGACTAAAGCAATAATTTATTTTCTCAATGTCTGTGC	1800
DB	1741	CTATTGGAGACACAAATATGCTGACTAAAGCAATAATTTATTTTCTCAATGTCTGTGC	1800
QY	1801	TAACTCAATGACTTAGAATGCTTTGCTATATTTTGGCTCTATGCTCAACACACTGGC	1860
DB	1801	TAACTCAATGACTTAGAATGCTTTGCTATATTTTGGCTCTATGCTCAACACACTGGC	1860
QY	1861	TTTTTTTTTAGCTCTTGAACAGCCAAACTGCTTCTGCTCAGGACCAAGATATTTTGGGA	1920
DB	1861	TTTTTTTTTAGCTCTTGAACAGCCAAACTGCTTCTGCTCAGGACCAAGATATTTTGGGA	1920
QY	1921	CTTCTCTTAAAGAAATCTATTTTCTTAAATCTTTATCTGGGTAACTTAGTTTATCCACA	1980
DB	1921	CTTCTCTTAAAGAAATCTATTTTCTTAAATCTTTATCTGGGTAACTTAGTTTATCCACA	1980
QY	1981	CTTCAGATCTGCGGCTTAAACCTCTTCTATAGAGGCTGTGATGACACTGTCTCTCTTC	2040
DB	1981	CTTCAGATCTGCGGCTTAAACCTCTTCTATAGAGGCTGTGATGACACTGTCTCTCTTC	2040
QY	2041	TCCACATACCTACACAGCACACATCTAGATAGATTAGAACCTCTCTGTTTTCTTTTCA	2100
DB	2041	TCCACATACCTACACAGCACACATCTAGATAGATTAGAACCTCTCTGTTTTCTTTTCA	2100
QY	2101	TACTTTTCTCTATCATGCTTCCCTCCATTAATAATTTTATTTATGTGTGTAATGTCTG	2160

Db 2101 TACTTTCTATCATGCTTCCCTCCATTATATATATTTTATATGTTGTAATGTTCTG 2160
Qy 2161 CCCAAGTCAGTTTCTCTCACTAACTATATAAATCCCTGTAAGCTGGGATCCTTCCAAATTT 2220
Db 2161 CCCAAGTCAGTTTCTCTCACTAACTATATAAATCCCTGTAAGCTGGGATCCTTCCAAATTT 2220
Qy 2221 GATCACCCTAGTACAGTAGGAACAACAGTAAGATTCATTTGTTATTTGTTGGAATGAAT 2280
Db 2221 GATCACCCTAGTACAGTAGGAACAACAGTAAGATTCATTTGTTATTTGTTGGAATGAAT 2280
Qy 2281 GAATGAATTTGTTGTTAGTAAAGCTTGGGGAAACCCAGGTGAGAGAGCCCTAGAAAGCA 2340
Db 2281 GAATGAATTTGTTGTTAGTAAAGCTTGGGGAAACCCAGGTGAGAGAGCCCTAGAAAGCA 2340
Qy 2341 GGTGGAATCCAAAGGCTAGATAGACTTAGTGTGTTACTCAAGAAAGGGTAGCCTGAAATATAA 2400
Db 2341 GGTGGAATCCAAAGGCTAGATAGACTTAGTGTGTTACTCAAGAAAGGGTAGCCTGAAATATAA 2400
Qy 2401 GGTTCAAATATAGTACAGAAATAGTCAAGATAGTGGGCAAGACAAGAGTCTCTGTTGCC 2460
Db 2401 GGTTCAAATATAGTACAGAAATAGTCAAGATAGTGGGCAAGACAAGAGTCTCTGTTGCC 2460
Qy 2461 GAATTCGATATCAAGCTTATCGATACCGTCCGACCTCGAGGGGGGCCCGGTACCCAAATTC 2520
Db 2461 GAATTCGATATCAAGCTTATCGATACCGTCCGACCTCGAGGGGGGCCCGGTACCCAAATTC 2520
Qy 2521 GCCCTATAGTGTGTTATTAACAATTCAGTGGCCGCTGTTTACAAACGTCGTGACTGGGA 2580
Db 2521 GCCCTATAGTGTGTTATTAACAATTCAGTGGCCGCTGTTTACAAACGTCGTGACTGGGA 2580
Qy 2581 AAACCTGGCGTTACCCAACTTAAT 2605
Db 2581 AAACCTGGCGTTACCCAACTTAAT 2605

RESULT 2

US-08-892-695-4
; Sequence 4, Application US/08892695A
; Patent No. 6808878
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W
; APPLICANT: Collins, Collin
; APPLICANT: Hwang, Soo In
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: GENES FROM THE 20Q13 AMPLICON AND THEIR USES
; FILE REFERENCE: 2500.124US3
; CURRENT APPLICATION NUMBER: US/08/892,695A
; EARLIER FILING DATE: 1997-07-15
; EARLIER FILING DATE: 1997-01-17
; EARLIER FILING DATE: 1997-01-17
; EARLIER FILING DATE: 1996-10-16
; EARLIER FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2605
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cc43
US-08-892-695-4

Query Match 100.0%; Score 2605; DB 4; Length 2605;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAAAGTCGAAATTAACCCCTCACTAAAGGAAACAAAGAGCTGGAGCTCCACCGGTTGGGG 60
|||||

Db 1 CAAAGTCGAAATTAACCCCTCACTAAAGGAAACAAAGAGCTGGAGCTCCACCGGTTGGCGG 60
Qy 61 CCGCTCTAGAACTAGTGGATCCCGGGCTGAGGAAATTCGACAGCTGGGCTACTATAC 120
|||||
Db 61 CCGCTCTAGAACTAGTGGATCCCGGGCTGAGGAAATTCGACAGCTGGGCTACTATAC 120
|||||
Qy 121 GATGGCGATGATTTGAGTGGCCGCTGGCAGTATGCTTCCACCCCTTCTTACGTTTACA 180
|||||
Db 121 GATGGCGATGATTTGAGTGGCCGCTGGCAGTATGCTTCCACCCCTTCTTACGTTTACA 180
|||||
Qy 181 ACCGAATGTGGACACTCGGCAAGCAGCTGCCCTCGTGTGCTCGTGTCTGCTGCTT 240
|||||
Db 181 ACCGAATGTGGACACTCGGCAAGCAGCTGCCCTCGTGTGCTCGTGTCTGCTGCTT 240
|||||
Qy 241 CTGCGCCCTGCAACAAAGTCCAGCATGACGCTGATGGAAGCTCAGAGAGAGCCCGCTCTT 300
|||||
Db 241 CTGCGCCCTGCAACAAAGTCCAGCATGACGCTGATGGAAGCTCAGAGAGAGCCCGCTCTT 300
|||||
Qy 301 CAAACAGTCAAGCTACAGCGAAAGCTTCTGTGAGTGCATCCAGATTTGTTATAGAGA 360
|||||
Db 301 CAAACAGTCAAGCTACAGCGAAAGCTTCTGTGAGTGCATCCAGATTTGTTATAGAGA 360
|||||
Qy 361 ACTGAGGAAGAAAGGAAACCTCGAGTGGTGGATGAAGCAAGTCCAGCTTCTGTATCAT 420
|||||
Db 361 ACTGAGGAAGAAAGGAAACCTCGAGTGGTGGATGAAGCAAGTCCAGCTTCTGTATCAT 420
|||||
Qy 421 GTGGCGGAGGCCAGAGAAATGGGGAAACTCATCTATCAGTGGGTTTCCAGAGAGTGCCCA 480
|||||
Db 421 GTGGCGGAGGCCAGAGAAATGGGGAAACTCATCTATCAGTGGGTTTCCAGAGAGTGCCCA 480
|||||
Qy 481 GAAACACTCCGCTTTTACCTGATGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAA 540
|||||
Db 481 GAAACACTCCGCTTTTACCTGATGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAA 540
|||||
Qy 541 GTTCCACGCGGCTGGATGAAGCCACTCTACTGCGGGCTCTGACAGGCCCTACACAGGAGCA 600
Db 541 GTTCCACGCGGCTGGATGAAGCCACTCTACTGCGGGCTCTGACAGGCCCTACACAGGAGCA 600
|||||
Qy 601 CAAAGCCGAGATCATCAGTGTGAGGAGGCGGAGGCGTCAAGTCTTCTAGCAGGAGCC 660
Db 601 CAAAGCCGAGATCATCAGTGTGAGGAGGCGGAGGCGTCAAGTCTTCTAGCAGGAGCC 660
|||||
Qy 661 TGTCTCTCTTACTTCTTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 720
Db 661 TGTCTCTCTTACTTCTTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 720
|||||
Qy 721 TGTCTCTCTTACTTCTTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 780
Db 721 TGTCTCTCTTACTTCTTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 780
|||||
Qy 781 CTGGGACCTGGGGATGGGTTTCTCACACCCCATATGTTCTGCTCTTCTTCTTCTTCTTCTTCTTCT 840
Db 781 CTGGGACCTGGGGATGGGTTTCTCACACCCCATATGTTCTGCTCTTCTTCTTCTTCTTCTTCTTCT 840
|||||
Qy 841 TGAAGCAACAGGAGAAATATGTTCTTCTGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900
Db 841 TGAAGCAACAGGAGAAATATGTTCTTCTGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900
|||||
Qy 901 TGTCTCTGAGCCAGGCTCTGTAACCTGACATTTATATGTTCTTCTTCTTCTTCTTCTTCTTCTTCT 960
Db 901 TGTCTCTGAGCCAGGCTCTGTAACCTGACATTTATATGTTCTTCTTCTTCTTCTTCTTCTTCTTCT 960
|||||
Qy 961 TACACATATGGGCTGACGACATGTTCTGCT 1020
Db 961 TACACATATGGGCTGACGACATGTTCTGCT 1020
|||||
Qy 1021 TGTTCCT 1080
Db 1021 TGTTCCT 1080
|||||
Qy 1081 TGCAGGAGCTCTTCCAGGCAATATGAAGTAGAGAGCCCAACGCGCTGCGAGTGAAGGT 1140
Db 1081 TGCAGGAGCTCTTCCAGGCAATATGAAGTAGAGAGCCCAACGCGCTGCGAGTGAAGGT 1140
|||||

1141 GTGGCCCAACCGATTATGATATATAAACTCAACTCCCAAAAAAAAAAAAAAAAAA 1200
Db
1141 GTGGCCCAACCGATTATGATATATAAACTCAACTCCCAAAAAAAAAAAAAAAAAA 1200
1201 CTGAGACTAGTTCTCTCTCTCGAAGAACTAGTCTCGAGTCTTTTCTTTTCTTTT 1260
Db
1201 CTGAGACTAGTTCTCTCTCTCGAAGAACTAGTCTCGAGTCTTTTCTTTTCTTTT 1260
1261 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1320
Db
1261 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1320
1321 ACTTTTCTCTACTTAATACTTTCTCAAGCTCTCAGAAGCCAGAGGGAAGAAAAAGAC 1380
Db
1321 ACTTTTCTCTACTTAATACTTTCTCAAGCTCTCAGAAGCCAGAGGGAAGAAAAAGAC 1380
1381 CATGAATCTTCTCTCCAGATTAAGTACACACTTTTGGAAAAACAGATTGGAAAACTTT 1440
Db
1441 CTGAAAAAGTTGACTGAAACTCCAAACCAACATGCCATATTTGTGATGTTGCTCATGAA 1500
Db
1441 CTGAAAAAGTTGACTGAAACTCCAAACCAACATGCCATATTTGTGATGTTGCTCATGAA 1500
1501 AATTGTTAAAAACCTGTTCTAGATAAGAAAGCTCTCAAGTTTGTACAGCCTACACAT 1560
Db
1501 AATTGTTAAAAACCTGTTCTAGATAAGAAAGCTCTCAAGTTTGTACAGCCTACACAT 1560
1561 AGTACAAGGGTCCCCATGATGATCTTCTGTAGGACGAAATAATGTAATTTTTTCAGTT 1620
Db
1561 AGTACAAGGGTCCCCATGATGATCTTCTGTAGGACGAAATAATGTAATTTTTTCAGTT 1620
1621 TCTGGTTTATAACTCTCTCGATCTCAGAGTTGACTGATTAAGAACCTACTCATGCAACA 1680
Db
1621 TCTGGTTTATAACTCTCTCGATCTCAGAGTTGACTGATTAAGAACCTACTCATGCAACA 1680
1681 GAGATAAAGCACTCATATTTTATAAATATATGACCAAACTATTTTGGAAATCTTAT 1740
Db
1681 GAGATAAAGCACTCATATTTTATAAATATATGACCAAACTATTTTGGAAATCTTAT 1740
1741 CTATTGGAGACACAATATGCTGGACTAAAGCAATAATTTATTTATCTCAATGCTGTGTC 1800
Db
1741 CTATTGGAGACACAATATGCTGGACTAAAGCAATAATTTATTTATCTCAATGCTGTGTC 1800
1801 TAACTCAATGACTTAGAATGCTTTGCTATATTTTGGCCTCTATGCTCAACCACTGGC 1860
Db
1801 TAACTCAATGACTTAGAATGCTTTGCTATATTTTGGCCTCTATGCTCAACCACTGGC 1860
1861 TTTCTTTTGTAGCTCTTGAACAAGCCAAAATGCTTCTGCTCAGGACCAAGATATTTGGGA 1920
Db
1861 TTTCTTTTGTAGCTCTTGAACAAGCCAAAATGCTTCTGCTCAGGACCAAGATATTTGGGA 1920
1921 CTTCTCTTAAAGAAATCTATTTCTTTAACTTTTATCTGGTAACTTAGTTTATCCAAACA 1980
Db
1921 CTTCTCTTAAAGAAATCTATTTCTTTAACTTTTATCTGGTAACTTAGTTTATCCAAACA 1980
1981 CTTGAGATCTGCGGTAAAACTCTTATAGAAGCCTGTATGACACTGCTCTCTCTTC 2040
Db
1981 CTTGAGATCTGCGGTAAAACTCTTATAGAAGCCTGTATGACACTGCTCTCTCTTC 2040
2041 TCCAACTACTCACCAGACACATGTAGACTAGATTAGAAGCTCTGTTTCTTTTCTTTTCA 2100
Db
2041 TCCAACTACTCACCAGACACATGTAGACTAGATTAGAAGCTCTGTTTCTTTTCTTTTCA 2100
2101 TACTTTTCTATCATGCTTCCCTCCATATAATAATTTTTATATGCTGTAATGCTG 2160
Db
2101 TACTTTTCTATCATGCTTCCCTCCATATAATAATTTTTATATGCTGTAATGCTG 2160
2161 CCCAAGTCAGTTTCTCCTCACTAAACTATAAAGCTCGTAAGCTGGATCCTTCCAAATTTT 2220
Db
2161 CCCAAGTCAGTTTCTCCTCACTAAACTATAAAGCTCGTAAGCTGGATCCTTCCAAATTTT 2220

2221 GATCACCACCTTAGTACAGTAGGAAACACAGTAAAGATTCAATTGGTATTTGTGGAATGAAT 2280
Db
2221 GATCACCACCTTAGTACAGTAGGAAACACAGTAAAGATTCAATTGGTATTTGTGGAATGAAT 2280
2281 GAATGAATTTGTTTCTAGTAAAGTCTGGGGGAAACCCAGGTGAGAAGAGCCTTAGAAAGCA 2340
Db
2281 GAATGAATTTGTTTCTAGTAAAGTCTGGGGGAAACCCAGGTGAGAAGAGCCTTAGAAAGCA 2340
2341 GGTGGAATCCAAAGCTAGATAGACTTAGTGTCTCAGAAGAGGTAGCTGAAATAAAA 2400
Db
2341 GGTGGAATCCAAAGCTAGATAGACTTAGTGTCTCAGAAGAGGTAGCTGAAATAAAA 2400
2401 GGTTCAAATATTAGTCAAGAAATAGTCAAGACATGGGCAAGACAGAGTGTCTGCTGCTGCC 2460
Db
2401 GGTTCAAATATTAGTCAAGAAATAGTCAAGACATGGGCAAGACAGAGTGTCTGCTGCTGCC 2460
2461 GAATTCGATATCAAGCTTATCGATACCGTCAAGCTCGAGGGGGGGCCGGTACCCAAATTC 2520
Db
2461 GAATTCGATATCAAGCTTATCGATACCGTCAAGCTCGAGGGGGGGCCGGTACCCAAATTC 2520
2521 GCCCTATAGTCAGTCGTATTACAAATTTCACTGGCGCTGTTTTTACAACTCGTGACTGGGA 2580
Db
2521 GCCCTATAGTCAGTCGTATTACAAATTTCACTGGCGCTGTTTTTACAACTCGTGACTGGGA 2580
2581 AAACCTGGCGCTTACCCCAACTTAAT 2605
Db
2581 AAACCTGGCGCTTACCCCAACTTAAT 2605

RESULT 3

US-09-620-312D-930
; Sequence 930, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 930
; LENGTH: 1243
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (172)..(702)
US-09-620-312D-930

Query Match 41.3%; Score 1076.2; DB 4; Length 1243;
Best Local Similarity 99.7%; Pred. No. 4e-247;

Matches 1078; Conservative 0; Mismatches 3; Indels 0; Gaps 0;				
QY	113	GCTACTACGATGGCGATGATTTCCGATGGCGGTGGCAGTATCGCTTCCACCCCTTTT	172	
DB	163	CGCTGACGATGGCGATGATTTCCGATGGCGGTGGCAGTATCGCTTCCACCCCTTTT	222	
QY	173	ACGTTACACCAATGTGACACTCGGCGAGAGCAGCTGGCGCCCTGGTGTCTGCTGTC	232	
DB	223	ACGTTACACCAATGTGACACTCGGCGAGAGCAGCTGGCGCCCTGGTGTCTGCTGTC	282	
QY	233	CTGTCTCTCTGCGCTGCAACAAAGTCCAGCATGACGCTGATGGAAGCTCAGGAGAC	292	
DB	283	CTGTCTCTCTGCGCTGCAACAAAGTCCAGCATGACGCTGATGGAAGCTCAGGAGAC	342	
QY	293	CGCTCTTTCAACACGTCAGCTACAGCGAAAGCTTCTGTGGAGTCGATCAGATTGTA	352	
DB	343	CGCTCTTTCAACACGTCAGCTACAGCGAAAGCTTCTGTGGAGTCGATCAGATTGTA	402	
QY	353	TTAGAGGAAGTGAAGAAAGGAACTCGAGTGGTGGATAAGAGCAAGTCCAGCTTC	412	
DB	403	TTAGAGGAAGTGAAGAAAGGAACTCGAGTGGTGGATAAGAGCAAGTCCAGCTTC	462	
QY	413	CTGATCATGTGCGGAGGCCAGAGAAATGGGGGAACTCATCTATCAGTGGGTTCACAG	472	
DB	463	CTGATCATGTGCGGAGGCCAGAGAAATGGGGGAACTCATCTATCAGTGGGTTCACAG	522	
QY	473	AGTGGCCAGAACAACTCCGTCTTTACCTGTATGAATGACTAAATGGGGAACACACAG	532	
DB	523	AGTGGCCAGAACAACTCCGTCTTTACCTGTATGAATGACTAAATGGGGAACACACAG	582	
QY	533	GATGAGGATTCACGGGCTGATGAGCACTCTATCGCGGCTCTGAGGCCCTACAG	592	
DB	583	GATGAGGATTCACGGGCTGATGAGCACTCTATCGCGGCTCTGAGGCCCTACAG	642	
QY	593	CAGGAGCAACAGCCGAGATCATCTGTCCGATGGCGGAGGCTCAAGTCTCTTAG	652	
DB	643	CAGGAGCAACAGCCGAGATCATCTGTCCGATGGCGGAGGCTCAAGTCTCTTAG	702	
QY	653	CAGGAGCTGTCTCCCTTTACTTTTACCTCCACCTTTCCAGGGGCTTCAAAGGAGAC	712	
DB	703	CAGGAGCTGTCTCCCTTTACTTTTACCTCCACCTTTCCAGGGGCTTCAAAGGAGAC	762	
QY	713	AGACCCAGTGTCCCGAAGACTGGATCTGTGACTCCACAGACTCAAAGGACTCCAGT	772	
DB	763	AGACCCAGTGTCCCGAAGACTGGATCTGTGACTCCACAGACTCAAAGGACTCCAGT	822	
QY	773	CCTGAAGCTGGGACTGGGATGGTCTCACCACCATATGTCTGCTCCCTCGGATAG	832	
DB	823	CCTGAAGCTGGGACTGGGATGGTCTCACCACCATATGTCTGCTCCCTCGGATAG	882	
QY	833	GGTGAGGCTGAAGCACACGAGGAGAAATATGTGCTTTCTCGCCCTACCTCTTTCCCA	892	
DB	883	GGTGAGGCTGAAGCACACGAGGAGAAATATGTGCTTTCTCGCCCTACCTCTTTCCCA	942	
QY	893	TCCTAGACTGTCTTTGAGCCAGGGCTGTGAAACCTGACACTTTATGTGTTTCAACATG	952	
DB	943	TCCTAGACTGTCTTTGAGCCAGGGCTGTGAAACCTGACACTTTATGTGTTTCAACATG	1002	
QY	953	TAAGTACATACACATGCGGCTGACGACATGCTTCTGCTCTCTCTCCACCCCT	1012	
DB	1003	TAAGTACATACACATGCGGCTGACGACATGCTTCTGCTCTCTCTCCACCCCT	1062	
QY	1013	TTAGTGTCTGTGCTCTCTCTGAGGATGGTCTGATCTCTCTAGGGGATGGGGAA	1072	
DB	1063	TTAGTGTCTGTGCTCTCTCTGAGGATGGTCTGATCTCTCTAGGGGATGGGGAA	1122	
QY	1073	GCCCTGTGTCAGGAGCCCTTCAGGCAATATGAAGATAGGAGGCCACCGGCTTGGCAG	1132	
DB	1123	GCCCTGTGTCAGGAGCCCTTCAGGCAATATGAAGATAGGAGGCCACCGGCTTGGCAG	1182	
QY	1133	TCAGAGGTGTGGCCCCACACCGATTTATGATATTAATCTCACTCCCAAAAAA	1192	
DB	1183	TCAGAGGTGTGGCCCCACACCGATTTATGATATTAATCTCACTCCCAAAAAA	1242	

QY 1193 A 1193
DB 1243 A 1243

RESULT 4
US-09-621-976-308
; Sequence 308, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 308
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 28..495
; NAME/KEY: sig peptide
; LOCATION: 28..156
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4
; OTHER INFORMATION: seq AAWCSLVLSFCRL/HK
US-09-621-976-308

Query Match 18.3%; Score 476; DB 4; Length 495;
Best Local Similarity 99.6%; Pred. No. 7.3e-104;
Matches 476; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY	109	CTGGGCTACTACGATGGCGATGAGTTTCGAGTGGCGGTGGCAGTATCGCTTCCACCCCTT	168	
DB	15	CTGGGCTACTACGATGGCGATGAGTTTCGAGTGGCGGTGGCAGTATCGCTTCCACCCCTT	74	
QY	169	CTTTACGTTTACAAACCGAATGTGGACACTCGGCGAGAAAGCAGCTGGCCGCTGTGTCTGCT	228	
DB	75	CTTTACGTTTACAAACCGAATGTGGACACTCGGCGAGAAAGCAGCTGGCCGCTGTGTCTGCT	134	
QY	229	GGTCTGTCTTCTGCGCGCTGCACAAACAGTCCAGCATGACGCTGATGGAAGCTCAGGA	288	
DB	135	GGTCTGTCTTCTGCGCGCTGCACAAACAGTCCAGCATGACGCTGATGGAAGCTCAGGA	194	
QY	289	GAGCCCGCTCTTCAACACGCTCAAGCTACAGCGAAAGCTTCTGTGGAGTCCGATCCAGAT	348	
DB	195	GAGCCCGCTCTTCAACACGCTCAAGCTACAGCGAAAGSTTCTGTGGAGTCCGATCCAGAT	254	
QY	349	TGTATTAGAGGAAGTGAAGAAAGGAAACCTCGAGTGGTGGATAAGAGCAAGTCCAG	408	
DB	255	TGTATTAGAGGAAGTGAAGAAAGGAAACCTCGAGTGGTGGATAAGAGCAAGTCCAG	314	
QY	409	CTTCTGTATCATGTGGCGGAGCCAGAAATGGGGGAACTCATCTATCAGTGGGTTC	468	
DB	315	CTTCTGTATCATGTGGCGGAGCCAGAAATGGGGGAACTCATCTATCAGTGGGTTC	374	
QY	469	CAGGATGGCCAGAACTCCGTCTTTTACCTGTATGAACCTGACTAATGGGGAAGACAC	528	
DB	375	CAGGATGGCCAGAACTCCGTCTTTTACCTGTATGAACCTGACTAATGGGGAAGACAC	434	
QY	529	AGAGGATGAGGAGTTCACAGGCGTGGATGAAGCCACTCTTACTGGGGCTCTGCAGGCC	586	
DB	435	AGAGGATGAGGAGTTCACAGGCGTGGATGAAGCCACTCTTACTCGGGCTCTGCAGGCC	492	

RESULT 5
US-09-621-976-307

```
; Sequence 307, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 307
; LENGTH: 487
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 38..487
; NAME/KEY: sig_peptide
; LOCATION: 38..166
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4
; OTHER INFORMATION: seq AAWCSLVLSFCRL/HK
US-09-621-976-307

Query Match          17.4%; Score 453; DB 4; Length 487;
Best Local Similarity 100.0%; Pred. No. 2.3e-98;
Matches 453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 CTGGGCTACTACGATGGCGATGAGTTTCGAGTGGCGGTGGCGATATCGTCTCCACCCCTT 168
DB 25 CTGGGCTACTACGATGGCGATGAGTTTCGAGTGGCGGTGGCGATATCGTCTCCACCCCTT 84

QY 169 CTTTACGTTTACAACCGAATGTGGACACTCGGCAGAACGAGCTGGCCGCTGGTCTCGCT 228
DB 85 CTTTACGTTTACAACCGAATGTGGACACTCGGCAGAACGAGCTGGCCGCTGGTCTCGCT 144

QY 229 GGTCTGTCTCTTCCGCGCTGCACAAACAGTCCAGCATGACGGTGATGGAGAGCTCAGGA 288
DB 145 GGTCTGTCTCTTCCGCGCTGCACAAACAGTCCAGCATGACGGTGATGGAGAGCTCAGGA 204

QY 289 GAGCCCGCTCTTCAACACGCTCAAGCTACAGCTACAGCG--AAGCTTCTCTGTGAGTCCAG 347
DB 85 CTTTACGTTTACAACCGAATGTGGACACTCGGCAGAACGAGCTGGCCGCTGGTCTCGCT 144

QY 229 GGTCTGTCTCTTCCGCGCTGCACAAACAGTCCAGCATGACGGTGATGGAGAGCTCAGGA 288
DB 145 GGTCTGTCTCTTCCGCGCTGCACAAACAGTCCAGCATGACGGTGATGGAGAGCTCAGGA 204

QY 349 TGTATTAGAGGAACGTGAGGAAGAAAGGAAACCTCGAGTGGTTGGATAAGAGCAAGTCCAG 408
DB 265 TGTATTAGAGGAACGTGAGGAAGAAAGGAAACCTCGAGTGGTTGGATAAGAGCAAGTCCAG 324

QY 409 CTTCTGTATCATGTGGCGGAGCCAGAGATGGGGGAACTCATCTATCATGTGGTTTC 468
DB 325 CTTCTGTATCATGTGGCGGAGCCAGAGATGGGGGAACTCATCTATCATGTGGTTTC 384

QY 469 CAGGAGTGCCAGAAACAACTCGCTCTTACCTCTGTATGAACCTGACTAATGGGGAAGAC 528
DB 385 CAGGAGTGCCAGAAACAACTCGCTCTTACCTCTGTATGAACCTGACTAATGGGGAAGAC 444

QY 529 AGAGGATGAGGAGTTTCCACGGGCTCGATGAAGC 561
DB 445 AGAGGATGAGGAGTTTCCACGGGCTCGATGAAGC 477

RESULT 6
US-09-621-976-1902
; Sequence 1902, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
```

```
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 1902
; LENGTH: 494
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 38..310
; NAME/KEY: sig_peptide
; LOCATION: 38..166
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4
; OTHER INFORMATION: seq AAWCSLVLSFCRL/HK
US-09-621-976-1902

Query Match          17.0%; Score 444; DB 4; Length 494;
Best Local Similarity 99.4%; Pred. No. 3.2e-96;
Matches 467; Conservative 0; Mismatches 0; Indels 3; Gaps 2;

QY 109 CTGGGCTACTACGATGGCGATGAGTTTCGAGTGGCGGTGGCGATATCGTCTCCACCCCTT 168
DB 25 CTGGGCTACTACGATGGCGATGAGTTTCGAGTGGCGGTGGCGATATCGTCTCCACCCCTT 84

QY 169 CTTTACGTTTACAACCGAATGTGGACACTCGGCAGAACGAGCTGGCCGCTGGTCTCGCT 228
DB 85 CTTTACGTTTACAACCGAATGTGGACACTCGGCAGAACGAGCTGGCCGCTGGTCTCGCT 144

QY 229 GGTCTGTCTCTTCCGCGCTGCACAAACAGTCCAGCATGACGGTGATGGAGAGCTCAGGA 288
DB 145 GGTCTGTCTCTTCCGCGCTGCACAAACAGTCCAGCATGACGGTGATGGAGAGCTCAGGA 204

QY 289 GAGCCCGCTCTTCAACACGCTCAAGCTACAGCTACAGCG--AAGCTTCTCTGTGAGTCCAG 347
DB 205 GAGCCCGCTCTTCAACACGCTCAAGCTACAGCGAAGAAAGCTTCTCTGTGAGTCCAG 264

QY 348 TTGTATTAGAGGAACGTGAGGAAGAAAGGAAACCTCGAGTGGTTGGATAAGAGCAAGTCCA 407
DB 265 TTGTATTAGAGGAACGTGAGGAAGAAAGGAAACCTCGAGTGGTTGGATAAGAGCAAGTCCA 324

QY 408 GTTCTCTGATCATGTGGCGGAGGCGCAGAAAGAAATGGGGG--AAACTCATCTATCATGTGGGT 465
DB 325 GTTCTCTGATCATGTGGCGGAGGCGCAGAAAGAAATGGGGGAAACTCATCTATCATGTGGGT 384

QY 466 TTCCAGGAGTGGCCAGAACAACTCGCTCTTTACCTGTATGAACCTGACTAATGGGGAAGA 525
DB 385 TTCCAGGAGTGGCCAGAACAACTCGCTCTTTACCTGTATGAACCTGACTAATGGGGAAGA 444

QY 526 CACAGAGGATGAGGAGTTCCACGGGCTGGATGAAGCCACTCTACTGCGGG 575
DB 445 CACAGAGGATGAGGAGTTCCACGGGCTGGATGAAGCCACTCTACTGCGGG 494

RESULT 7
US-09-621-976-15507
; Sequence 15507, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15507
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapiens
```

;	FEATURE:																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
---	----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Db 300 CTGATCATGTGGCGGANGCCA-AAGAAATGGGGAAACTCATCTACGTGGTTTCCA-G 357
Qy 473 AGTGGCCGAACAACTCCGTCTTTTACCCCTGTATGAACAGCTAA 516
Db 358 ANTGCCANAACTCCGTCTTTTACCCCTGTATGAACACTAA 401

RESULT 10
US-09-621-976-15508
; Sequence 15508, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15508
; LENGTH: 383
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: 183..184
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-15508

Query Match 12.6%; Score 328.2; DB 4; Length 383;
Best Local Similarity 96.8%; Pred. No. 1.4e-68;
Matches 359; Conservative 6; Mismatches 3; Indels 3; Gaps 3;

Qy 109 CTGGCTACTACGATGGCGATGAGTTTCGAGTGGCCGTGGCAGTATCGCTTCCCAACCCCTT 168
Db 15 CTGGCTACTACGATGGCGATGAGTTTCGAGTGGCCGTGGCAGTATCGCTTCCCAACCCCTT 74
Qy 169 CTTTACGTT-ACACCGAATGGACACTCGGCAGAACGAGCTGGCCGCTGGTCTCGC 227
Db 75 CTTTAMGRTAAACACGAAATGGACACTCGGCAGAACGAGCTGGCCGCTGGTCTCGC 134
Qy 228 TGGTCTCTCTCTTCGCGCCCTGCACAAACAGTCAGCATGACGCTGATGGAAGCTCAGG 287
Db 135 TGGTCTCTCTCTTCGCGCCCTGCACAAACAGTCAGCATGACGCTVHNGAAGCTCAGG 194
Qy 288 AGAGCCCGCTCTTCAACAAACGTCACAGCTACAGCGAAAGCTTCTCTGGAGTCGATCCAGA 347
Db 195 AGAGCCCGCTCTTCAACAAACGTCACAGCTACAGCGAAA-STTCTGTGGAGTCGATCCAGA 253
Qy 348 TTGATTAGAGAACTGAGGAGAAAGGAACTCGAGTGGTGGATAGAGCAAGTCCA 407
Db 254 TTGATTAGAGAACTGAGGAGAAAGGAACTCGAGTGGTGGATAGAGCAAGTCCA 313
Qy 408 GCTTCTCATCATGTGGCGAGGCCAGAGAATGGGGAACTCATCTATCAGTGGGTTT 467
Db 314 GCTTCTCATCATGTGGCGAGG-CAGAGAATGGGGAACTCATCTATCAGTGGGTTT 372
Qy 468 CCAGGAGTGGC 478
Db 373 CCAGGAGTGGC 383

RESULT 11
US-09-621-976-15509
; Sequence 15509, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15509
; LENGTH: 360
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15509

Query Match 11.9%; Score 309; DB 4; Length 360;
Best Local Similarity 96.4%; Pred. No. 5.2e-64;
Matches 325; Conservative 10; Mismatches 0; Indels 2; Gaps 2;

Qy 109 CTGGCTACTACGATGGCGATGAGTTTCGAGTGGCCGTGGCAGTATCGCTTCCCAACCCCTT 168
Db 26 CTGGCTACTACGATGGCGATGAGTTTCGAGTGGCCGTGGCAGTATCGCTTCCCAACCCCTT 85
Qy 169 CTTTACGTTTACAACCGAATGGACACTCGGCAGAACGAGCTGGCCGCTGGTCTCGCT 228
Db 86 CTTTACGTTTACAACCGAATGGACACTCGGCAGAACGAGCTGGCCGCTGGTCTCGCT 145
Qy 229 GGTCTGTCTTCTTCGCGCCCTGCACAAACAGTCACGATGACGGTGTGGAAGCTCAGGA 288
Db 146 GGTCTGTCTTCTTCGCGCCCTGCACAAACAGTCACGATGACGGTGTGGAAGCTCAGGA 205
Qy 289 GAGCCCGCTCTTCAACAAACGTCAGCTACAGCGAAAGCTTCTCTGGAGTGCATCCAGAT 348
Db 206 GAGCCCGCTCTTCAACAAACGTCAGCTACAGCGAAA-STTCTGTGGAGTCGATCCAGAT 264
Qy 349 TGTATTAGAGAACTGAGGAGAAAGGAACTCGAGTGGTGGATAAGAGCAAGTCCAG 408
Db 265 TGTATTAGAGAACTGAGGAGAAAGGAACTCGARTGGTGGATAAGAGCAAGTCCAG 324
Qy 409 CTTCTGTATCATGTGGCGAGGCCAGAGAAATGGGGG 445
Db 325 CTTCTGTATCATGTGGCGARG-CAGAGAATGGGGG 360

RESULT 12
US-09-702-705-1380/c
; Sequence 1380, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121-478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1380
; LENGTH: 269
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-1380

Query Match 10.3%; Score 269; DB 4; Length 269;
Best Local Similarity 100.0%; Pred. No. 1.6e-54;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 209 CTGGCCGCTCGTGTCTCGCTCGCTGTCTTCTTGGCCGCTGCACAAACAGTCCAGCATG 268


```

; TYPE: DNA
; ORGANISM: Homo sapien
US-09-671-325-1380

Query Match      10.3%; Score 269; DB 4; Length 269;
Best Local Similarity 100.0%; Pred. No. 1.6e-54;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 209 CTGGCCGCTGGTCTCGTCTGCTGCTCTTCTTCTGCGCCTGCACAAACAGTCCAGCATG 268
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 269 CTGGCCGCTGGTCTCGTCTGCTGCTCTTCTTCTGCGCCTGCACAAACAGTCCAGCATG 210
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 269 ACGGTGATGGAGCTCAGGAGAGCCGCTCTTCAACAAACGTCAAGCTACAGCGAAAGCTT 328
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 209 ACGGTGATGGAGCTCAGGAGAGCCGCTCTTCAACAAACGTCAAGCTACAGCGAAAGCTT 150
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 329 CCTGTGGAGTCGATCCAGATTGATTAGAGGAACTGAGGAGAAAGGAAACCTCGAGTGG 388
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 149 CCTGTGGAGTCGATCCAGATTGATTAGAGGAACTGAGGAGAAAGGAAACCTCGAGTGG 90
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 389 TTGGATAGAGCAAGTCCAGCTTCTGATCATGTGGCGGAGCCAGAGATGGGGGAAA 448
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 89 TTGGATAGAGCAAGTCCAGCTTCTGATCATGTGGCGGAGCCAGAGATGGGGGAAA 30
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 449 CTCATCTATCAGTGGGTTTCCAGGAGTGG 477
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 29 CTCATCTATCAGTGGGTTTCCAGGAGTGG 1
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Search completed: June 29, 2005, 05:27:54
Job time : 299.707 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2005, 01:06:08 ; Search time 1019 Seconds
(without alignments)
15978.923 Million cell updates/sec

Title: US-08-731-499-4
Perfect score: 2605
Sequence: 1 CRAAGTCGAATTAAACCTC.....CTGGCGTTACCACTTAAT 2605

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 6067389 seqs, 3125258755 residues

Total number of hits satisfying chosen parameters: 12134778

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2:	/cgn2_6/ptodata/1/pubpna/PT_NEW_PUB.seq.*
3:	/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6:	/cgn2_6/ptodata/1/pubpna/PTUS_PUBCOMB.seq.*
7:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9:	/cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10:	/cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11:	/cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13:	/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14:	/cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15:	/cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16:	/cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17:	/cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18:	/cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19:	/cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20:	/cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21:	/cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
23:	/cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24:	/cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26:	/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2605	100.0	2605	8	US-08-731-499-4
2	1076.2	41.3	1243	15	Sequence 4, Appl
3	1076.2	41.3	1243	17	Sequence 930, App
4	1073	41.2	1087	10	Sequence 930, App
5	603.6	23.2	634	11	Sequence 25, Appl
6	515.8	19.8	559	11	Sequence 727, App
7	464.4	17.8	507	11	Sequence 636, App
					Sequence 533, App

8	395	15.2	609	11	US-09-969-034-610	Sequence 610, App
9	353.6	13.6	545	17	US-10-115-482-13	Sequence 13, Appl
10	317.4	12.2	349	11	US-09-969-034-665	Sequence 665, App
11	269	10.3	269	9	US-09-998-598-2364	Sequence 2264, App
12	269	10.3	269	9	US-09-998-598-2400	Sequence 2400, App
13	269	10.3	269	9	US-09-736-457-1380	Sequence 1380, App
14	269	10.3	269	9	US-09-902-941-1380	Sequence 1380, App
15	269	10.3	269	9	US-09-849-628-1380	Sequence 1380, App
16	269	10.3	269	9	US-09-796-692-6790	Sequence 6790, App
17	269	10.3	269	14	US-10-017-754-1380	Sequence 1380, App
18	269	10.3	269	14	US-10-040-862-6790	Sequence 6790, App
19	269	10.3	269	16	US-10-113-872-1380	Sequence 1380, App
20	269	10.3	269	17	US-10-283-017-1380	Sequence 1380, App
21	269	10.3	269	17	US-10-057-475B-6790	Sequence 6790, App
22	269	10.3	269	17	US-10-154-884B-6790	Sequence 6790, App
23	269	10.3	269	19	US-10-764-324-6790	Sequence 6790, App
24	267.4	10.3	269	9	US-09-796-692-8647	Sequence 8647, App
25	267.4	10.3	269	14	US-10-040-862-8647	Sequence 8647, App
26	267.4	10.3	269	17	US-10-057-475B-8647	Sequence 8647, App
27	267.4	10.3	269	17	US-10-154-884B-8647	Sequence 8647, App
28	267.4	10.3	269	19	US-10-764-324-8647	Sequence 8647, App
29	260	10.0	260	9	US-09-998-598-218	Sequence 218, App
30	260	10.0	260	9	US-09-796-692-3665	Sequence 3665, App
31	260	10.0	260	14	US-10-040-862-3665	Sequence 3665, App
32	260	10.0	260	17	US-10-057-475B-3665	Sequence 3665, App
33	260	10.0	260	17	US-10-154-884B-3665	Sequence 3665, App
34	260	10.0	260	19	US-10-764-324-3665	Sequence 3665, App
35	257	9.9	260	13	US-10-033-528-1797	Sequence 1797, App
36	257	9.9	260	16	US-10-099-926-1797	Sequence 1797, App
37	227	8.7	227	9	US-09-796-692-7056	Sequence 7056, App
38	227	8.7	227	14	US-10-040-862-7056	Sequence 7056, App
39	227	8.7	227	17	US-10-057-475B-7056	Sequence 7056, App
40	227	8.7	227	17	US-10-154-884B-7056	Sequence 7056, App
41	227	8.7	227	19	US-10-764-324-7056	Sequence 7056, App
42	227	8.7	241	21	US-10-926-683-611	Sequence 611, App
43	219.2	8.4	469	20	US-10-357-930-37999	Sequence 37999, A
44	218.2	8.4	335	20	US-10-357-930-25157	Sequence 25157, A
45	169.4	6.5	217	10	US-09-918-995-19060	Sequence 19060, A

ALIGNMENTS

RESULT 1

US-08-731-499-4
; Sequence 4, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION:
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOMBEL, David
; APPLICANT: ROMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,499
; FILING DATE: 16-OCT-1996
; CLASSIFICATION: 435

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/680,395
; FILING DATE: 15-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 23070-068910
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2605 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..2605 /note= "cDNA clone cc43 of 4 kb
; OTHER INFORMATION:
; OTHER INFORMATION: transcript"
US-08-731-499-4

Query Match      100.0%; Score 2605; DB 8; Length 2605;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCTCGAAATTAAACCTCTACTAAAGGGAACAAAGCTGGAGCTCCACCGCGTGGCGG 60
DB 1 CAAGCTCGAAATTAAACCTCTACTAAAGGGAACAAAGCTGGAGCTCCACCGCGTGGCGG 60

QY 61 CGGCTCTAGAACTAGTGATCCCGCGGCTGCAGGAATTCGGCAGAGCTGGCTACTAC 120
DB 61 CGGCTCTAGAACTAGTGATCCCGCGGCTGCAGGAATTCGGCAGAGCTGGCTACTAC 120

QY 121 GATGCGGATGAGTTTCGAGTGGCCCTGGCAGTATCGCTTCCACCCCTCTTTTACGTTTACA 180
DB 121 GATGCGGATGAGTTTCGAGTGGCCCTGGCAGTATCGCTTCCACCCCTCTTTTACGTTTACA 180

QY 181 ACCGAATGTGACACTCGGCGAGAAGCAGCTGGCCGCTGGTGTCTCGCTGGTCTGTCCTT 240
DB 181 ACCGAATGTGACACTCGGCGAGAAGCAGCTGGCCGCTGGTGTCTCGCTGGTCTGTCCTT 240

QY 241 CTGCCGCTGCACAAACAGTCCAGATGACGCTGATGGAAGCTCAGGAGAGCCGCTCTT 300
DB 241 CTGCCGCTGCACAAACAGTCCAGATGACGCTGATGGAAGCTCAGGAGAGCCGCTCTT 300

QY 301 CAACAACTCAAGCTACAGCGAAAGCTTCTCTGGAGTCCGATCCAGATTGTATTAGAGGA 360
DB 301 CAACAACTCAAGCTACAGCGAAAGCTTCTCTGGAGTCCGATCCAGATTGTATTAGAGGA 360

QY 361 ACTGAGGAAGAAAGGAACCTCGAGTGGTTGATTAAGAGCAAGTCCAGCTTCTGATCAT 420
DB 361 ACTGAGGAAGAAAGGAACCTCGAGTGGTTGATTAAGAGCAAGTCCAGCTTCTGATCAT 420

QY 421 GTGGCGGAGCCAGAGGAATGGGGAACTCATCTATCAGTGGGTTTCCAGAGTGGCCA 480
DB 421 GTGGCGGAGCCAGAGGAATGGGGAACTCATCTATCAGTGGGTTTCCAGAGTGGCCA 480

QY 481 GAACAACTCCGCTCTTTACCTGTATGAACCTGACTAATGGGGAAGACACAGAGGATGAGGA 540
DB 481 GAACAACTCCGCTCTTTACCTGTATGAACCTGACTAATGGGGAAGACACAGAGGATGAGGA 540

QY 541 GTTCCACGGGCTGGATGAAGCCACTCTACTGCGGCTCTGAGGCCCTTACAGCAGGAGA 600
DB 541 GTTCCACGGGCTGGATGAAGCCACTCTACTGCGGCTCTGAGGCCCTTACAGCAGGAGA 600

QY 601 CAAGGCCAGATCATCACTGTCCAGGATGGCCGAGCGCTCAAGTCTTCTTACGAGGACC 660
DB 601 CAAGGCCAGATCATCACTGTCCAGGATGGCCGAGCGCTCAAGTCTTCTTACGAGGACC 660

QY 661 TGTCTCCCTTTTACTTCTTACCTCCACCTTTTCCAGGGCTTTTCAAAAGGAGACAGACCCAG 720

```

```

DB 661 TGTCTCCCTTTTACTTCTTTTACCTCCACCTTTCCAGGGCTTTCAAAAGGAGACAGACCCAG 720
QY 721 TGTCCCCCAAAAGACTGGGATCTGTGACTCCACAGACTCAAAAGGACTCCAGTCTCTGAAGG 780
DB 721 TGTCCCCCAAAAGACTGGGATCTGTGACTCCACAGACTCAAAAGGACTCCAGTCTCTGAAGG 780
QY 781 CTGGGACCTGGGGATGGGTTTCTCACACCCCATATGTCTGCTCCCTTGGATAGGGTGAAGC 840
DB 781 CTGGGACCTGGGGATGGGTTTCTCACACCCCATATGTCTGCTCCCTTGGATAGGGTGAAGC 840
QY 841 TGAAGCACACGAGGAGAAATATGTCTTCTCTGCCCCCTACTCTCTTCCCCTATCCATAGAC 900
DB 841 TGAAGCACACGAGGAGAAATATGTCTTCTCTGCCCCCTACTCTCTTCCCCTATCCATAGAC 900
QY 901 TGTCTTGGAGCCAGGGTCTGTAAACCTGACACTTTATATGTGTTCACACATGTAAGTACA 960
DB 901 TGTCTTGGAGCCAGGGTCTGTAAACCTGACACTTTATATGTGTTCACACATGTAAGTACA 960
QY 961 TACACACATGCGCTGACGACATGCTTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
DB 961 TACACACATGCGCTGACGACATGCTTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
QY 1021 TGTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1080
DB 1021 TGTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1080
QY 1081 TGCAGGAGCCCTTCCAGGCAATATGAAGATAGGAGGCCCAAGGCTTGGCAGTGAAGGT 1140
DB 1081 TGCAGGAGCCCTTCCAGGCAATATGAAGATAGGAGGCCCAAGGCTTGGCAGTGAAGGT 1140
QY 1141 GTGGCCCCCAGACCGGATTTATGATATTAATAATCTCAACTCCCAAAAAAAGGAGGAGG 1200
DB 1141 GTGGCCCCCAGACCGGATTTATGATATTAATAATCTCAACTCCCAAAAAAAGGAGGAGG 1200
QY 1201 CTGAGACTAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
DB 1201 CTGAGACTAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
QY 1261 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1320
DB 1261 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1320
QY 1321 ACTTTCTCTACTTAAATACTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380
DB 1321 ACTTTCTCTACTTAAATACTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380
QY 1381 CATGAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1440
DB 1381 CATGAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1440
QY 1441 CTGAAAAAGTGTGACTGAACTCCAAACCAACATGCCATATTTGTTGATGTTGCTCATGAA 1500
DB 1441 CTGAAAAAGTGTGACTGAACTCCAAACCAACATGCCATATTTGTTGATGTTGCTCATGAA 1500
QY 1501 AATTGTTAAAAACCTGTTCTAGATAAAGAAAGCTCTCAAGTTTCTTGTACAGCTTACACAT 1560
DB 1501 AATTGTTAAAAACCTGTTCTAGATAAAGAAAGCTCTCAAGTTTCTTGTACAGCTTACACAT 1560
QY 1561 AGTACAGGGTCCCTATGATGATTTCTCTGTAGGACGAAATAATGTAATTTTTCAGTT 1620
DB 1561 AGTACAGGGTCCCTATGATGATTTCTCTGTAGGACGAAATAATGTAATTTTTCAGTT 1620
QY 1621 TCTGGTTTATAACTCTCTCGATCTCAGAGTTGACGATTAAACACCTTACTCATCAACA 1680
DB 1621 TCTGGTTTATAACTCTCTCGATCTCAGAGTTGACGATTAAACACCTTACTCATCAACA 1680
QY 1681 CAGATAAAGCACTCATATTTTATAAATATATGAGCAAACTATTTTGGAAATCTTAT 1740
DB 1681 CAGATAAAGCACTCATATTTTATAAATATATGAGCAAACTATTTTGGAAATCTTAT 1740
QY 1741 CTATTGGAGACACATATGCTGGACTAAAGCAATTAATTTTATTTTCTCAATGTCGTGTC 1800

```

Db 1741 CTATTGGACACAAATATCTGGACTAAAGCAATAATATTTATTTCTCAATGTCTGTGC 1800
Qy 1801 TAACTCAATAGCTTAGAATGCTTTGCTATATTTTGCCTCTATGCTCAACACACTGGC 1860
Db 1801 TAACTCAATAGCTTAGAATGCTTTGCTATATTTTGCCTCTATGCTCAACACACTGGC 1860
Qy 1861 TTTCTTTTAGCTTTGGAACAGCAACAGCTTCTGCTCTAGGACAGATATTTTGGGA 1920
Db 1861 TTTCTTTTAGCTTTGGAACAGCAACAGCTTCTGCTCTAGGACAGATATTTTGGGA 1920
Qy 1921 CTTCTCTTAAGAAATCTATTTCTTAATTTCTTTATCTGGTAACCTTAGTTTATCCACA 1980
Db 1921 CTTCTCTTAAGAAATCTATTTCTTAATTTCTTTATCTGGTAACCTTAGTTTATCCACA 1980
Qy 1981 CTTGAGATCTCGCGTAAACCTTTCTTTATAGAAGCTGTGATGACACTGCTCTCTTC 2040
Db 1981 CTTGAGATCTCGCGTAAACCTTTCTTTATAGAAGCTGTGATGACACTGCTCTCTTC 2040
Qy 2041 TCCACATACCTCACAGGACACATGTAGACTAGATTAGAACCTCTGTTTCTTTTCA 2100
Db 2041 TCCACATACCTCACAGGACACATGTAGACTAGATTAGAACCTCTGTTTCTTTTCA 2100
Qy 2101 TACTTTTCTCTATCATGCTTCTCCCTCCATTAATATTTTATTTATGTTGTAATGTCG 2160
Db 2101 TACTTTTCTCTATCATGCTTCTCCCTCCATTAATATTTTATTTATGTTGTAATGTCG 2160
Qy 2161 CCCAAGTCAGTTTCTCTAACTATAAACTCCGTAAAGCTGGGATCTTCCAAATTT 2220
Db 2161 CCCAAGTCAGTTTCTCTAACTATAAACTCCGTAAAGCTGGGATCTTCCAAATTT 2220
Qy 2221 GATCACCATTAGTACAGTAAAGATTAAGATTTCAATTTGTTGTTGTAATGAAT 2280
Db 2221 GATCACCATTAGTACAGTAAAGATTAAGATTTCAATTTGTTGTTGTAATGAAT 2280
Qy 2281 GAATGAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2340
Db 2281 GAATGAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2340
Qy 2341 GGTGGAATCCAAAGGCTAGATAGACTTTAGTTGTTTACTCAAGAAAGGTTAGCTTGA 2400
Db 2341 GGTGGAATCCAAAGGCTAGATAGACTTTAGTTGTTTACTCAAGAAAGGTTAGCTTGA 2400
Qy 2401 GGTTCAAATATAGTCAAGAAATAGTCAAGAAATAGTCAAGAAATAGTCAAGAAATAGT 2460
Db 2401 GGTTCAAATATAGTCAAGAAATAGTCAAGAAATAGTCAAGAAATAGTCAAGAAATAGT 2460
Qy 2461 GAATTCGATATCAAGCTTATCGATACCGTCCGACCTCGAGGGGGCCCGGTACCAATTC 2520
Db 2461 GAATTCGATATCAAGCTTATCGATACCGTCCGACCTCGAGGGGGCCCGGTACCAATTC 2520
Qy 2521 GCCCTATAGTGTGATTTTACAAATTTCACTGGCCGCTGTTTATCAACGTCGTGAGTGGGA 2580
Db 2521 GCCCTATAGTGTGATTTTACAAATTTCACTGGCCGCTGTTTATCAACGTCGTGAGTGGGA 2580
Qy 2581 AAACCTGGCGTTTACCAACTTAAT 2605
Db 2581 AAACCTGGCGTTTACCAACTTAAT 2605

RESULT 2

US-10-037-270-930
; Sequence 930, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenchua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.

; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: NO. US20030104529A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP28
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 930
; LENGTH: 1243
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (172)..(702)
US-10-037-270-930

Query Match 41.3%; Score 1076.2; DB 15; Length 1243;
Best Local Similarity 99.7%; Pred. No. 4.2e-237;
Matches 1078; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 113 GCTACTACGATGGCGATGAGTTTCGAGTGGCGGTGCGAGTATCGCTTCCACCCCTTTCTTT 172
Db 163 GCGTCGACGATGGCGATGAGTTTCGAGTGGCGGTGCGAGTATCGCTTCCACCCCTTTCTTT 222
Qy 173 AGTTTAAACCAATGTGGACACTCGGAGAGAGAGCTGGCGCCCTGGTGTCTGCTGGTTC 232
Db 223 AGTTTAAACCAATGTGGACACTCGGAGAGAGAGCTGGCGCCCTGGTGTCTGCTGGTTC 282
Qy 233 CTGTCCTCTTCTGCGGCTGCACAAACAGTCCAGCATGACGCTGATGAAAGCTCAGAGAGC 292
Db 283 CTGTCCTCTTCTGCGGCTGCACAAACAGTCCAGCATGACGCTGATGAAAGCTCAGAGAGC 342
Qy 293 CGCTCTTTCAACAACAGCTCAAGCTACAGCGAAAGCTTCTGTGGAGTTCGATCCAGATTGTA 352
Db 343 CGCTCTTTCAACAACAGCTCAAGCTACAGCGAAAGCTTCTGTGGAGTTCGATCCAGATTGTA 402
Qy 353 TTAGAGGAACCTGAGGAAGAGGAACTCTGAGTGGTTGGATAGAGCAAGTCCAGCTTC 412
Db 403 TTAGAGGAACCTGAGGAAGAGGAACTCTGAGTGGTTGGATAGAGCAAGTCCAGCTTC 462
Qy 413 CTGATCATGTGGCGGAGCCAGAAAGTGGGGGAACTCATCTATCAGTGGGTTTCCAGG 472
Db 463 CTGATCATGTGGCGGAGCCAGAAAGTGGGGGAACTCATCTATCAGTGGGTTTCCAGG 522
Qy 473 AGTGGCCAGAACAACTCCGTCTTTTACCTCTGTATGAATCACTAAATGGGGAAGACACAGAG 532
Db 523 AGTGGCCAGAACAACTCCGTCTTTTACCTCTGTATGAATCACTAAATGGGGAAGACACAGAG 582
Qy 533 GATGAGGATTTCCAGGGCTGGATGAAGCACTCTATCTGGGGCTCTGAGGGCCCTACAG 592
Db 583 GATGAGGATTTCCAGGGCTGGATGAAGCACTCTATCTGGGGCTCTGAGGGCCCTACAG 642
Qy 593 CAGGAGCAAGGCGGAGATCATCTGTACGGATGGCGGGCTCAAGTTCTTCTAG 652
Db 643 CAGGAGCAAGGCGGAGATCATCTGTACGGATGGCGGGCTCAAGTTCTTCTAG 702
Qy 653 CAGGAGCACTGTCTCTCTTTTACTTTTACCTCCACCTTTTCCAGGGCTTTCAAAGAGAGAC 712
Db 703 CAGGAGCACTGTCTCTCTTTTACTTTTACTCTCCACCTTTTCCAGGGCTTTCAAAGAGAGAC 762
Qy 713 AGACCCAGTGTCCCCCAAGACTGGATCTGTGATCTCCACAGACTCAAAAGGACTCCAGT 772

Db	1243 A 1243				
RESULT 4					
US-09-890-688-25					
; Sequence 25, Application US/09890688					
; Publication No. US20030144475A1					
; GENERAL INFORMATION:					
; APPLICANT: Seishi KATO					
; APPLICANT: Chikashi EGUCHI					
; APPLICANT: Mihoro SAEKI					
; TITLE OF INVENTION: Human Proteins and cDNAs thereof					
; FILE REFERENCE: 2001-1102A/WMC/00653					
; CURRENT APPLICATION NUMBER: US/09/890,688					
; CURRENT FILING DATE: 2001-08-06					
; PRIOR APPLICATION NUMBER: JP 11-346863					
; PRIOR FILING DATE: 1999-12-06					
; PRIOR APPLICATION NUMBER: JP 11-34684					
; PRIOR FILING DATE: 1999-12-06					
; PRIOR APPLICATION NUMBER: JP 2000-31062					
; PRIOR FILING DATE: 2000-02-08					
; PRIOR APPLICATION NUMBER: JP 2000-34091					
; PRIOR FILING DATE: 2000-02-10					
; PRIOR APPLICATION NUMBER: JP 2000-34090					
; PRIOR FILING DATE: 2000-02-10					
; PRIOR APPLICATION NUMBER: JP 2000-35829					
; PRIOR FILING DATE: 2000-02-14					
; PRIOR APPLICATION NUMBER: JP 2000-35899					
; PRIOR FILING DATE: 2000-02-14					
; PRIOR APPLICATION NUMBER: JP 2000-71161					
; PRIOR FILING DATE: 2000-03-14					
; PRIOR APPLICATION NUMBER: JP 2000-160851					
; PRIOR FILING DATE: 2000-05-30					
; NUMBER OF SEQ ID NOS: 160					
; SOFTWARE: PatentIn Ver. 2.1					
; SEQ ID NO 25					
; LENGTH: 1087					
; TYPE: DNA					
; ORGANISM: Homo sapiens					
; FEATURE:					
; NAME/KEY: CDS					
; LOCATION: (28)..(558)					
US-09-890-688-25					
Query Match 41.2%; Score 1073; DB 10; Length 1087;					
Best Local Similarity 100.0%; Pred. No. 2.le-236;					
Matches 1073; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	109	CTGGGCTACTACGATGGCGATGAGTTTCGAGTGGCCGTGGCAGTATCGTTCCACCCTT	168		
Db	15	CTGGGCTACTACGATGGCGATGAGTTTCGAGTGGCCGTGGCAGTATCGTTCCACCCTT	74		
Qy	169	CTTTAGTTTACACCGAATGTGCACACTCGGCAGAACGCTGCCCGCTGGTGCTCGCT	228		
Db	75	CTTTAGTTTACACCGAATGTGCACACTCGGCAGAACGCTGCCCGCTGGTGCTCGCT	134		
Qy	229	GGTCCTGTCTCTTCTGCGGCTGCACAAACAGTCCAGCATGACGGTGTGAAGCTCAGGA	288		
Db	135	GGTCCTGTCTCTTCTGCGGCTGCACAAACAGTCCAGCATGACGGTGTGAAGCTCAGGA	194		
Qy	289	GAGCCCGCTCTTCAACAACAGTCAAGCTACAGCGAAAGCTTCTGTGGAGTCGATCCAGAT	348		
Db	195	GAGCCCGCTCTTCAACAACAGTCAAGCTACAGCGAAAGCTTCTGTGGAGTCGATCCAGAT	254		
Qy	349	TGTATTAGGAAGTAAGGAAGAAGGAAAGGAAAGTTCGAGTGGTGTGATAGCAAGTCCAG	408		
Db	255	TGTATTAGGAAGTAAGGAAGAAGGAAAGGAAAGTTCGAGTGGTGTGATAGCAAGTCCAG	314		
Qy	409	CTTCTCTGATCATGTGCGGAGCCAGAGAAATGGGGAAACTCATCTATCATGTGGTTTTTC	468		
Db	315	CTTCTCTGATCATGTGCGGAGCCAGAGAAATGGGGAAACTCATCTATCATGTGGTTTTTC	374		
Qy	469	CAGGAGTGGCCAGAACAACTCGCTCTTTTACCTGTATGAAGTCACTAATGGGGAGACAC	528		

Db	375	CAGGAGTGGCCAGAACAACTCGCTCTTTTACCTGTATGAAGTCACTAATGGGGAGACAC	434
Qy	529	AGAGGATGAGGAGTTTCCACGGGCTGGATGAAGCCACTCTACTCGGGGCTCTGAGGCGCT	588
Db	435	AGAGGATGAGGAGTTTCCACGGGCTGGATGAAGCCACTCTACTCGGGGCTCTGAGGCGCT	494
Qy	589	ACAGCAGGAGCACAAAGGCCGAGATCATCTGTGAGGATGGCCGAGGCGTCAAGTTCTT	648
Db	495	ACAGCAGGAGCACAAAGGCCGAGATCATCTGTGAGGATGGCCGAGGCGTCAAGTTCTT	554
Qy	649	CTAGCAGGAGCTGTCTCCCTTTACTTCTTACCTCCCACCTTTCCAGGGCTTTCAAAGG	708
Db	555	CTAGCAGGAGCTGTCTCCCTTTACTTCTTACCTCCCACCTTTCCAGGGCTTTCAAAGG	614
Qy	709	AGACAGACCCAGTGTCCCCCAAAGACTGGATCTGTGACTCCACAGACTCAAAGGACTC	768
Db	615	AGACAGACCCAGTGTCCCCCAAAGACTGGATCTGTGACTCCACAGACTCAAAGGACTC	674
Qy	769	CAGTCTGAAGGCTGGGACCTGGGGATGGGTTTCTCACCCCCATATGTCTGCTCCCTTG	828
Db	675	CAGTCTGAAGGCTGGGACCTGGGGATGGGTTTCTCACCCCCATATGTCTGCTCCCTTG	734
Qy	829	ATAGGTTGAGGCTGAAGCAGCAGGAGAAATATGTCTTCTCGCCCTACCTCCTT	888
Db	735	ATAGGTTGAGGCTGAAGCAGCAGGAGAAATATGTCTTCTCGCCCTACCTCCTT	794
Qy	889	CCATCTTAGACTGTCTTGGAGCAGGCTGTGAAAACCTGACACTTTATATGTTTCAACA	948
Db	795	CCATCTTAGACTGTCTTGGAGCAGGCTGTGAAAACCTGACACTTTATATGTTTCAACA	854
Qy	949	CATGTAAGTACATACACATGCGCTGCAGCAGCATGCTTCTCTCTCTCTCTCCAC	1008
Db	855	CATGTAAGTACATACACATGCGCTGCAGCAGCATGCTTCTCTCTCTCTCCAC	914
Qy	1009	CCCTTTAGCTGTGTGCT	1068
Db	915	CCCTTTAGCTGTGTGCT	974
Qy	1069	GGAGCCCTGGCTGCAGCAGCCTTCCAGGCAATATGAAGATAGGAGGCCACCGGGCTG	1128
Db	975	GGAGCCCTGGCTGCAGCAGCCTTCCAGGCAATATGAAGATAGGAGGCCACCGGGCTG	1034
Qy	1129	GCAGTGAGAGTGTGGCCCAACCGGATTTATGATATATAAATCTCAACTCCC	1181
Db	1035	GCAGTGAGAGTGTGGCCCAACCGGATTTATGATATATAAATCTCAACTCCC	1087
RESULT 5			
US-09-969-034-727/c			
; Sequence 727, Application US/09969034			
; Publication No. US20040110668A1			
; GENERAL INFORMATION:			
; APPLICANT: Burgess, Christopher C.			
; APPLICANT: Astle, Jon H.			
; APPLICANT: Carroll, Eddie III			
; APPLICANT: Catino, Theodore J.			
; APPLICANT: Dwivedi, Poornima			
; APPLICANT: Molino, Gary A.			
; APPLICANT: Thiagalingam, Arunthathi			
; APPLICANT: Lewis, Marcia E.			
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially			
; FILE REFERENCE: 1657/1032			
; CURRENT APPLICATION NUMBER: US/09/969,034			
; CURRENT FILING DATE: 2001-10-02			
; PRIOR APPLICATION NUMBER: 60/237,271			
; PRIOR FILING DATE: 2000-02-10			
; NUMBER OF SEQ ID NOS: 4494			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 727			
; LENGTH: 634			
; TYPE: DNA			

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 577_
OTHER INFORMATION: n = A,T,C or G
US-09-969-034-727

Query Match 23.2%; Score 603.6; DB 11; Length 634;
Best Local Similarity 98.7%; Pred. No. 1.9e-128;
Matches 629; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

QY 321 GAAAGCTTCCTGTGGAGTCGATCCAGATTGTATTAGAGAACTGAGGAAGAAAGGAAC 380
DB 634 GAAAGCTTCCTGTGGAGTCGATCCAGATTGTATTAGAGAA--TTAAGAAGAAGGAACN 577
QY 381 TCGAGTGGTTGGATAAGAGCAAGTCAGCTTCCTGATCATGTGGCGGAGCCAGAGAAT 440
DB 576 TCGAGTGGTTGGATAAGAGCAAGTCAGCTTCCTGATCATGTGGCGGAGCCAGAGAAT 517
QY 441 GGGGAAACTCATCTATCAGTGGGTTTCCAGAGTGGCCAGAACTCCGTTCTTTACCC 500
DB 516 GGGGAAACTCATCTATCAGTGGGTTTCCAGAGTGG--CAGAACTCCGTTCTTTACCC 458
QY 501 TGTATGAACACTGACTAATGGGGAAGACACAGAGGATGAGGAGTTCCACGGGCTGGATGAAG 560
DB 457 TGTATGAACACTGACTAATGGGGAAGACACAGAGGATGAGGAGTTCCACGGGCTGGATGAAG 398
QY 561 CCACCTACTCGGGCTCTCGAGGCCCTACAGCAGGAGCAGAGGCCAGATCATCACTG 620
DB 397 CCACCTACTCGGGCTCTCGAGGCCCTACAGCAGGAGCAGAGGCCAGATCATCACTG 338
QY 621 TCAGGAGTGGCGAGGCGTCAAGTTCTTCTAGCAGGACCTGTCTCCCTTTACTTCTTAC 680
DB 337 TCAGGAGTGGCGAGGCGTCAAGTTCTTCTAGCAGGACCTGTCTCCCTTTACTTCTTAC 278
QY 681 CTCACACCTTTCCAGGGCTTTCAAAGGAGACAGACCAGTGTCCTCCCAAGACTGGATC 740
DB 277 CTCACACCTTTCCAGGGCTTTCAAAGGAGACAGACCAGTGTCCTCCCAAGACTGGATC 218
QY 741 TGTGACTCCACAGACTCAAAGGACTCCAGTCTGAGGCTGGAGCTGGGGATGGGTT 800
DB 217 TGTGACTCCACAGACTCAAAGGACTCCAGTCTGAGGCTGGAGCTGGGGATGGGTT 158
QY 801 TCTCACACCCATATGTCTGCTCCCTTGGATAGGTTGAGGCTGAAGCACACAGGAGAAAT 860
DB 157 TCTCACACCCATATGTCTGCTCCCTTGGATAGGTTGAGGCTGAAGCACACAGGAGAAAT 98
QY 861 ATGTGCTTTCTTCGCCCCTACCTCTTTCCCATCTAGACTGTCTTGGCCAGGGTCTG 920
DB 97 ATGTGCTTTCTTCGCCCCTACCTCTTTCCCATCTAGACTGTCTTGGCCAGGGTCTG 38
QY 921 TAAACCTGACACTTTATATGTGTTCCACATGTAAGT 957
DB 37 TAAACCTGGCACTTTATATGTGTTCCACATGTAAGT 1

RESULT 6
US-09-969-034-636
; Sequence 636, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Pooruima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034

CURRENT FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: 60/237,271
PRIOR FILING DATE: 2000-02-10
NUMBER OF SEQ ID NOS: 4494
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 636
LENGTH: 559
TYPE: DNA
ORGANISM: Homo sapiens
US-09-969-034-636

Query Match 19.8%; Score 515.8; DB 11; Length 559;
Best Local Similarity 99.6%; Pred. No. 3e-108;
Matches 517; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 109 CTGGGCTACTACGATGGCGATGAGTTTCGAGTGGCCGTCGAGTATCCGTTCCACCCCTT 168
DB 16 CTGGGCTACTACGATGGCGATGAGTTTCGAGTGGCCGTCGAGTATCCGTTCCACCCCTT 75
QY 169 CTTTACGTTTACAACCGAATGTGGACACTCCGCGAGAGCAGCTGGCCGCTGGTGTCTCGCT 228
DB 76 CTTTACGTTTACAACCGAATGTGGACACTCCGCGAGAGCAGCTGGCCGCTGGTGTCTCGCT 135
QY 229 GGTCTCTCTCTTCTCCGCTTCACAAACAGTCCAGCATGACGGTATGGAAGCTCAGGA 288
DB 136 GGTCTCTCTCTTCTCCGCTTCACAAACAGTCCAGCATGACGGTATGGAAGCTCAGGA 195
QY 289 GAGCCGCTCTTCAACAACGTCAGCTACAGGAAAGCTTCTGTGGAGTCGATCCAGAT 348
DB 196 GAGCCGCTCTTCAACAACGTCAGCTACAGGAAAGCTTCTGTGGAGTCGATCCAGAT 255
QY 349 TGTATTACAGGAACTGAGGAAGAAAGGAACTCCAGTGGTTGGATAAGAGCAAGTCCAG 408
DB 256 TGTATTACAGGAACTGAGGAAGAAAGGAACTCCAGTGGTTGGATAAGAGCAAGTCCAG 315
QY 409 CTTCTCTGATCATGTGGCGAGGCCAGAAAGAAATGCGGGAACCTCATCTATCATGTGGTTTC 468
DB 316 CTTCTCTGATCATGTGGCGAGGCCAGAAAGAAATGCGGGAACCTCATCTATCATGTGGTTTC 375
QY 469 CAGGAGTGGCCAGAACTCCGCTTTTACCTGTATCACTGACTATATGGGGAAGACAC 528
DB 376 CAGGAGTGGCCAGAACTCCGCTTTTACCTGTATCACTGACTATATGGGGAAGACAC 435
QY 529 AGAGGATCAGGAGTTCCACGGCTGGATGAAGCCACTTACTGCGGGCTCTGACGGCCCT 588
DB 436 AGAGGATCAGGAGTTCCACGGCTGGATGAAGCCACTTACTGCGGGCTCTGACGGCCCT 495
QY 589 ACAGCAGGACCAAGCCGAGATCATCTGTGACGGA 627
DB 496 ACAGCAGGACCAAGCCGAGATCATCTACTGTGACGGA 534

RESULT 7
US-09-969-034-533/c
; Sequence 533, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Pooruima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 533
; LENGTH: 507
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 327, 340, 357, 368, 469, 498
; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-533

Query Match 17.8%; Score 464.4; DB 11; Length 507;
Best Local Similarity 97.0%; Pred. No. 1.9e-96;
Matches 491; Conservative 0; Mismatches 12; Indels 3; Gaps 2;

QY 455 TATCAGTGGGTTTCCAGGA--GTGGCCAGAACCACTCCGTCTTTACCCCTGTATGAAC-TG 511
DB 506 TATCAGTGGGTTTCCAGGAAGTGGCCAGAACCACTCNGTCTTTTACCCCTGTATGAAC 447

QY 512 ACTAATGGGGAGACACAGAGGATGAGGAGTTCCACGGGCTGGATGAAGCACTCTACTG 571
DB 446 ACTAATGGGGAGACACAGAGGATGAGGAGTTCCACGGGCTGGATGAAGCACTTTACTG 387

QY 572 CGGGCTCTGCAGGCCCTACAGCAGGAGCAAGGCCGAGATCATCACTGTGAGCGATGCG 631
DB 386 CGGGCTCTGCAGGCCCTTANAGCAGGAGCANAGGCCGAGATCATCACTGTGAGCGATG 327

QY 632 CGAGGCGTCAAGTCTTCTAGCAGGACCTGTCTCCCTTTTACTTCTTACCTTCCACCTTT 691
DB 326 CGAGGCGTCAAGTCTTCTAGCAGGACCTGTCTCCCTTTTACTTCTTACCTTCCACCTTT 267

QY 692 CGAGGCGTTCAGAGGAGACAGACCCAGTGTCCCCCAAGACTGATCTGTGACTCCAC 751
DB 266 CGAGGCGTTCAGAGGAGACAGACCCAGTGTCCCCCAAGACTGATCTGTGACTCCAC 207

QY 752 CAGACTCAAAAGGACTCCAGTCTGAAGGCTGGGACTGGGATGGTGTCTTCCACACCC 811
DB 206 CAGACTCAAAAGGACTCCAGTCTGAAGGCTGGGACTGGGATGGTGTCTTCCACACCC 147

QY 812 ATATGTCTCTCCCTGGATAGGCTGAGGCTGAAGCACCAGGAGAGAAATATGTCTTCT 871
DB 146 ATATGTCTCTCCCTGGATAGGCTGAGGCTGAAGCACCAGGAGAGAAATATGTCTTCT 87

QY 872 CTCGCCCTACCTCTTTTCCATCTCTAGACTGTCTTGGCCAGGCTCTGTAACTGACA 931
DB 86 CTCGCCCTACCTCTTTTCCATCTCTAGACTGTCTTGGCCAGGCTCTGTAACTGACA 27

QY 932 CTTTATATGTGTTCCACATGTAAGT 957
DB 26 CTTTATATGTGTTCCACATGTAAGT 1

RESULT 8
US-09-969-034-610
; Sequence 610, Application US/09969034
; Publication No. US2004011068A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poorima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 610
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 468, 478, 491, 501, 505, 574, 576, 577, 604
; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-610

Query Match 15.2%; Score 395; DB 11; Length 609;
Best Local Similarity 96.5%; Pred. No. 2e-80;
Matches 447; Conservative 0; Mismatches 10; Indels 6; Gaps 4;

QY 112 GGCTACTACGATGGCGGATGATTTTCAGTGGCGGTGGCAGTATCGCTTCCACCCCTTCTT 171
DB 5 GGGGACTACGATGGGTGATGAGTTTCGAGTGGCGGTGGCAGTATCGCTTCCACCCCTTCTT 64

QY 172 TACGTTAACCAACCGAATGTGGACACTCGGCAGAACGAGCTGGCGCTGTGCTCGTGTGT 231
DB 65 TACGTTAACCAACCGAATGTGGACACTCGGCAGAACGAGCTGGCGCTGTGCTCGTGTGT 124

QY 232 CTGTGCTCTTCTCCGCTCGCACAAACAGTCCAGCATGACGCTGATGGAAGCTCCAGAGAG 291
DB 125 CTGTGCTCTTCTCCGCTCGCACAAACAGTCCAGCATGACGCTGATGGAAGCTCCAGAGAG 184

QY 292 CCGGCTCTTTCACCAACAGTCAAGCTACAGCGAAAGCTTCTGTGGAGTCCGATCCAGATTGT 351
DB 185 CCGGCTCTTTCACCAACAGTCAAGCTACAGCGAAAGCTTCTGTGGAGTCCGATCCAGATTGT 244

QY 352 ATTAGAGGAACGTGAGGAAGAAAGGAAACCTCGAGTGGTTGGATAAGAGCAAGTCCAGCTT 411
DB 245 ATTAGAGGAACGTGAGGAAGAAAGGAAACCTCGAGTGGTTGGATAAGAAACAAGTCCAGCTT 304

QY 412 CCGTATCATGTGGCGGAGCCAGAAAGATGGGGGAAACTCATCTATCAGTGGGTTCCAG 471
DB 305 CCGTATCATGTGGCGGAG--GCCAGAAAGATGGGGGAAACTCATCTATCAGTGG--TTTCAG 361

QY 472 GAGTGGCCAGAAACAACCTCGTCTTTTACCCTGTATGAACCTGACTTAATGGGGAAGACACAGA 531
DB 362 GAGTGGC--AGAACACTCGTCTTTTACCCTGTATGAACCTGACTTAATGGGGAAGACACAGA 419

QY 532 GGATGAGGAGTTCCACGGGCTGGATGAAGCCACTCTACTGCGG 574
DB 420 GGATGAGGAGTTCCACGGGCTGGATGA--CCACTCTACTGCGG 461

RESULT 9
US-10-115-482-13
; Sequence 13, Application US/10115482
; Publication No. US20030212257A1
; GENERAL INFORMATION:
; APPLICANT: Spytex, et al.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
; TITLE OF INVENTION: AND METHODS
; TITLE OF INVENTION: OF USING THE SAME
; FILE REFERENCE: 21404-322D
; CURRENT APPLICATION NUMBER: US/10/115,482
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/281,086
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/285,325

RESULT 10
US-09-969-034-665
; Sequence 665, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poorima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; TITLE OF INVENTION: Expressed in Cancer Tissue
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 665
; LENGTH: 349
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 233, 274, 278, 321
; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-665

Query Match 13.6%; Score 353.6; DB 17; Length 545;
Best Local Similarity 84.6%; Pred. No. 6.4e-71;
Matches 435; Conservative 0; Mismatches 69; Indels 10; Gaps 3;

QY 109 CTGGCTACTACGATGGCGATGAGTTTCAGAGTGGCCGTGGCAGTATCGCTTCCACCCCTT 168
Db 8 CTGGCTATTATGATGGTGCAATTTCAAGTGGTTGGCAGTATCGCTTCCGCC--- 64
QY 169 CTTTACGTTTACAACCGAATGTGACACTCGGCAGAGCAGCTGGCCGCTGGTGTCTCGCT 228
Db 65 CTTTACATTACAGCTGAAGTGGCCACTTGGCAGAGCAGCTGGCCACTGGTGTGTGT 124
QY 229 GGTCTGTCTTCTCCGCTGCACAAACAGTCCAGCATGACGCTGATGGAAGCTCAGGA 288
Db 125 GGTCTGTCTTCTGCTGCTGCACAGACAGTCAAGCATGATGGTTATGGATGCTTCAGGA 184
QY 289 GAGCCCGCTCTTCAACAACGTCACGCTACAGGAAGCTTCTGTGGAGTGCATCCAGAT 348
Db 185 GATCTGTCTTCTCAGCAATCAGCT--GTGGAGCTTCTGTGGATCATCCAGT 241
QY 349 TGTTATTAGGAACCTGAGGAAGAAAGGAACTTCAGTGGTTGGATGAAGAGCAAGTCCAG 408
Db 242 TGTTATTAGGAACCTGAGGAAGAAATGGAACTTACAGTGGCTGGATGAAGAGCAAGTCTAG 301
QY 409 CTTCTGTATCATGTGGCGGAGCCAGAGAATGGGGGAACTCATCTATCATGTGGTTTC 468
Db 302 TTTCTTAATCATGTGGCGGAGCCAGAGAATGGGGGAACTCATCTATCATGTGGGTCTC 361
QY 469 CAGGATGGCCAGAACAACTCGTCTTTACCTCTGATGAACCTGACTTAATGGGGAAGACAC 528
Db 362 CAGGATGGCCAGAACAACTCGTACTTATGCTGATGAGCTGACCAATGGGGAAGACAT 421
QY 529 AGAGGATGAGGAGTTCCACGGCTGGATCAAGCCACTCTACTCGGGCTCTGCAGGCCCT 588
Db 422 AGAGAATGAGGTGTTCCACGGACTAAAGGAGGC---CTTCTGTGGCTCTGCAGGCCCT 477
QY 589 ACAGCAGGACCAAGCGCGAGATCATCTGTC 622
Db 478 TCAGTAGGAACATAGGCTGAGATCATCACCATC 511

Query Match 12.2%; Score 317.4; DB 11; Length 349;
Best Local Similarity 98.5%; Pred. No. 1.1e-62;
Matches 318; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 637 CGTCAAGTCTTCTAGCAGGACCTGTCTCCCTTTTACTTCTTACCTCCACCTTTCCAGG 696
Db 1 CGTCAAGTCTTCTAGCAGGACCTGTCTCCCTTTTACTTCTTACCTCCACCTTTCCAGG 60
QY 697 GCTTTCAAAGGAGACAGACCCAGTGTCCCCCAAGACTGGATCTGTGACTCCACGAC 756
Db 61 GCTTTCAAAGGAGACAGACCCAGTGTCCCCCAAGACTGGATCTGTGACTCCACGAC 120
QY 757 TCAAAAGACTCCAGTCTCTGAAGCTGGACCTGGGATGGTTTCTCACACCCCATATG 816
Db 121 TCAAAAGACTCCAGTCTCTGAAGCTGGGACCTGGGATGGTTTCTCACACCCCATATG 180
QY 817 TCTGTCCCTTCGATAGGTGAGGCTGAAGCACCAGGAGAAAAATATGTGCTTCTCTCGC 876
Db 181 TCTGTCCCTTCGATAGGTGAGGCTGAAGCACCAGGAGAAAAATATGTGCTTCTCTCGC 240
QY 877 CTTACCTCTTTTCCCATCTTAGACTGTCTTTGAGCCAGGCTGTGTAACCTGACACTTTA 936
Db 241 CTTACCTCTTTTCCCATCTTAGACTGTCTTTGANCANGGTCTGTAAACCTGACACTTTA 300
QY 937 TATGTGTTACACATGTAAGTAC 959
Db 301 TATGTGTTACACATGTAAGNCC 323

RESULT 11
US-09-998-598-2264/c
; Sequence 2264, Application US/09998598
; Patent No. US20020150922A1
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Chenault, Ruth A.
; APPLICANT: Mesgher, Madelein Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 15:06:07 ; Search time 506.03 Seconds
(without alignments)
17687.639 Million cell updates/sec

Title: US-08-731-499-4

Perfect score: 2605

Sequence: 1 CAAGCTCGAATTAACCTC.....CTGGCGTTACCACTTAAT 2605

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1041	40.0	1041	3	CR620674 full-leng
2	1040	39.9	1060	3	CR611687 full-leng
3	1016	39.0	1016	3	CR615047 full-leng
4	965.4	37.1	1021	1	AL561088 AL561088
5	947.2	36.4	1079	5	BX357325 BX357325
6	923	35.4	1039	5	BM927839 AGENCOURT
7	914.8	35.1	971	1	AL582492 AL582492
8	903	34.7	1005	4	BM554015 AGENCOURT
9	889	34.1	982	6	CA487570 AGENCOURT
10	879.2	33.8	972	5	BX383031 BX383031
11	868	33.3	918	5	BQ672154 AGENCOURT
12	859	33.0	936	7	COS81180 ILLUMIGEN
13	854.6	32.8	1049	5	BM919274 AGENCOURT
14	844.8	32.4	925	4	BG675348 AGENCOURT
15	843.8	32.4	990	4	BM555365 AGENCOURT
16	838	32.2	959	5	BX357450 AGENCOURT
17	831.6	31.9	1046	4	BM810698 AGENCOURT
18	830.4	31.9	901	4	BM018273 AGENCOURT
19	826.4	31.7	961	5	BQ882132 AGENCOURT
20	822.8	31.6	897	6	CA455057 AGENCOURT
21	814.6	31.3	859	4	BG674662 AGENCOURT
22	813.6	31.2	919	4	BG742758 AGENCOURT
23	811.4	31.1	957	7	COT26143 ILLUMIGEN
24	811.2	31.1	943	5	BQ673129 AGENCOURT

25	811	31.1	909	7	COS82238	COS82238 ILLUMIGEN
26	807	31.0	908	5	BU540421	BU540421 AGENCOURT
27	804.8	30.9	877	5	BU168204	BU168204 AGENCOURT
28	799.4	30.7	898	6	CA489194	CA489194 AGENCOURT
29	798.8	30.7	981	4	BM474026	BM474026 AGENCOURT
30	796.6	30.6	877	4	BI550414	BI550414 AGENCOURT
31	790.8	30.4	842	1	AL528911	AL528911 AGENCOURT
32	790.6	30.3	928	6	CA489833	CA489833 AGENCOURT
33	778.6	29.9	923	5	BQ646055	BQ646055 AGENCOURT
34	776.8	29.8	779	1	AL528912	AL528912 AGENCOURT
35	775.2	29.8	863	5	BU183716	BU183716 AGENCOURT
36	775.2	29.8	877	5	BU153711	BU153711 AGENCOURT
37	772.4	29.7	911	6	CA487561	CA487561 AGENCOURT
38	771.2	29.6	1027	5	BQ053239	BQ053239 AGENCOURT
39	770.4	29.6	1074	3	AK003715	AK003715 Mus muscu
40	769.2	29.5	1057	4	BM476382	BM476382 AGENCOURT
41	767.8	29.5	872	6	CD388777	CD388777 AGENCOURT
42	766.8	29.4	1067	3	AK002500	AK002500 Mus muscu
43	751	28.8	974	4	BM561840	BM561840 AGENCOURT
44	750.8	28.8	829	4	BG281239	BG281239 602402192
45	745.4	28.6	803	5	BQ773588	BQ773588 UI-H-PEO-

ALIGNMENTS

RESULT 1	CR620674	1041 bp	mRNA	linear	HTC 21-JUL-2004
LOCUS	full-length cDNA clone CS0DD004YG22 of Neuroblastoma Cot				
DEFINITION	50-normalized of Homo sapiens (human).				
ACCESSION	CR620674.1	GI:50501481			
VERSION	HTC; CNSLT cDNA.				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 1041)				
AUTHORS	Li, W.B., Gruber, C., Jessee, J., and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished				
REMARK	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paraday Avenue				
REFERENCE	2 (bases 1 to 1041)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
COMMENT	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life technologies, a division of invitrogen.				
FEATURES	Location/Qualifiers				
source	1..1041				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="CS0DD004YG22"				
	/tissue_type="Neuroblastoma Cot 50-normalized"				
	/plasmid="pCMVSPORT_6"				

Query Match	40.0%;	Score 1041;	DB 3;	Length 1041;
Best Local Similarity	100.0%;	Pred. No. 1.4e-217;		
Matches 1041;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	129	TCAGTTTCGAGTGGCGGTGGCAGTATCGTTCCACCCCTCTTACGTTACACCGAATG	188	
Db	1	TGAGTTTCGAGTGGCGGTGGCAGTATCGTTCCACCCCTCTTACGTTACACCGAATG	60	


```
QY 589 ACAGCAGGAGCACAAGGCGGAGATCATCTGTGAGCGATGCGCGAGGCGTCAAGTTCTT 648
Db 501 ACAGCAGGAGCACAAGGCGGAGATCATCTGTGAGCGATGCGCGAGGCGTCAAGTTCTT 560
QY 649 CTAGCAGGAGCCTGTCTCCCTTTACTTCTTACTCTTACTCTCCACCTTTCCAGGCGTTTCAAAAGG 708
Db 561 CTAGCAGGAGCCTGTCTCCCTTTACTTCTTACTCTCCACCTTTCCAGGCGTTTCAAAAGG 620
QY 709 AGACAGACCCAGTGTCCCCAAAGACTGGATCTGTGACTTCCACAGACTCAAAAGGACTC 768
Db 621 AGACAGACCCAGTGTCCCCAAAGACTGGATCTGTGACTTCCACAGACTCAAAAGGACTC 680
QY 769 CAGTCTGAAGGCTGGGACCTGGGATGGGTTTCTCTACACCCCATATGTCTGTCCTTGG 828
Db 681 CAGTCTGAAGGCTGGGACCTGGGATGGGTTTCTCTACACCCCATATGTCTGTCCTTGG 740
QY 829 ATAGGCTGAGGCTGAAGCAGCAGGAGAAATATGTCTTCTTCGCGCTACCTCCCTT 888
Db 741 ATAGGCTGAGGCTGAAGCAGCAGGAGAAATATGTCTTCTTCTTCGCGCTACCTCCCTT 800
QY 889 CCCATCCTAGACTGTCTTGAGCCAGGCTGTGTAAACCTGACACTTTATATGTGTTCACA 948
Db 801 CCCATCCTAGACTGTCTTGAGCCAGGCTGTGTAAACCTGACACTTTATATGTGTTCACA 860
QY 949 CATGTAAGTACATACACATGCGGCTGAGCAGACATGTCTGTCTCTCTCTCCCTCCAC 1008
Db 861 CATGTAAGTACATACACATGCGGCTGAGCAGACATGTCTGTCTCTCTCTCTCCCTCCAC 920
QY 1009 CCCTTTAGCTGTGTCTCTCTCTCTCAGGCTGTGTGTGATCTCTCTAGGGGATGG 1068
Db 921 CCCTTTAGCTGTGTCTCTCTCTCTCAGGCTGTGTGTGATCTCTCTAGGGGATGG 980
QY 1069 GGAAGCCCTGGCTGAGGAGGCTTCCAGGCAATATGAAGATAGGAGGCCACCGGCGCTG 1128
Db 981 GGAAGCCCTGGCTGAGGAGGCTTCCAGGCAATATGAAGATAGGAGGCCACCGGCGCTG 1040
QY 1129 GCAGTGAGAGGTGTGGCCCC 1148
Db 1041 GCAGTGAGAGGTGTGGCCCC 1060

RESULT 3
LOCUS CR615047
DEFINITION full-length cDNA clone CS0DC007YC08 of Neuroblastoma Cot
ACCESSION CR615047
VERSION CR615047.1 GI:50495854
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1016)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
2 (bases 1 to 1016)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen
FEATURES
Location/Qualifiers
1..1016
source
```

ORIGIN

```
Query Match 39.0%; Score 1016; DB 3; Length 1016;
Best Local Similarity 100.0%; Pred. No. 4.2e-212;
Matches 1016; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 TGAGTTTCAGTGGCGGTGGCAGATATCGTTCGCCACCTTCTTTAGCTTACAAACGAATG 188
Db 1 TGAGTTTCAGTGGCGGTGGCAGATATCGTTCGCCACCTTCTTTAGCTTACAAACGAATG 60
QY 189 TGGACACTCGCAGCAAGCAGCTGGCGCTGTGTCTCGTGTCTGTCTCTCTCTCTCTCTCT 248
Db 61 TGGACACTCGCAGCAAGCAGCTGGCGCTGTGTCTCGTGTCTGTCTCTCTCTCTCTCTCT 120
QY 249 TGCACAAACAGTCCAGCATGACGGTGTGGAAGCTTCAGGAGAGCCCGCTCTTTCAACAACG 308
Db 121 TGCACAAACAGTCCAGCATGACGGTGTGGAAGCTTCAGGAGAGCCCGCTCTTTCAACAACG 180
QY 309 TCAAGCTACAGGAAAGCTTCTCTGTGTGAGTGTGATCCAGATTGTATTAGAGGAATGAGGA 368
Db 181 TCAAGCTACAGGAAAGCTTCTCTGTGTGAGTGTGATCCAGATTGTATTAGAGGAATGAGGA 240
QY 369 AGAAGGGAACCTCGAGTGGTTGGATAGAGCAAGTCCAGCTTCTGTGATCATGTGCGCGA 428
Db 241 AGAAGGGAACCTCGAGTGGTTGGATAGAGCAAGTCCAGCTTCTGTGATCATGTGCGCGA 300
QY 429 GGCCAGAAAGATGGGGGAAACTCATCTATCATGTGGTTTCCAGGAGTGGCCAGAAACAAT 488
Db 301 GGCCAGAAAGATGGGGGAAACTCATCTATCATGTGGTTTCCAGGAGTGGCCAGAAACAAT 360
QY 489 CCGTCTTTACCTGTATGAACTGACTAATGGGGAAGACACAGAGGATGAGGAGTTCCACG 548
Db 361 CCGTCTTTACCTGTATGAACTGACTAATGGGGAAGACACAGAGGATGAGGAGTTCCACG 420
QY 549 GCCTGGATGAAGCCACTCTACTTGGGGCTCTGCAGGCGCTTCAAGAGGAGCAACAGGCG 608
Db 421 GCCTGGATGAAGCCACTCTACTTGGGGCTCTGCAGGCGCTTCAAGAGGAGCAACAGGCG 480
QY 609 AGATCATCACTGTTCAGCGATGGCGAGGCGTCAAGTTCTTCTTAGCAGGAGACCTGTCTCC 668
Db 481 AGATCATCACTGTTCAGCGATGGCGAGGCGTCAAGTTCTTCTTAGCAGGAGACCTGTCTCC 540
QY 669 TTTACTTCTTACCTCCACCTTTTCCAGGCGCTTCAAAAGGAGACAGACCCAGTGTCCCC 728
Db 541 TTTACTTCTTACCTCCACCTTTTCCAGGCGCTTCAAAAGGAGACAGACCCAGTGTCCCC 600
QY 729 AAAGACTGGATCTGTGACTCCACAGACTCAAAAGGACTCCAGTCTTCAAGGCTGGGACC 788
Db 601 AAAGACTGGATCTGTGACTCCACAGACTCAAAAGGACTCCAGTCTTCAAGGCTGGGACC 660
QY 789 TGGGATGGGTTTCTCACACCCCATATGTCTGTCTCTTGGATAGGGTGAAGGCTGAAGCAC 848
Db 661 TGGGATGGGTTTCTCACACCCCATATGTCTGTCTCTTGGATAGGGTGAAGGCTGAAGCAC 720
QY 849 CAGGAGAAATATGTGCTTCTTGGCGCTTCTGTCTCTCTTCCATCTCTAGACTGTCTTTG 908
Db 721 CAGGAGAAATATGTGCTTCTTGGCGCTTCTGTCTCTCTTCCATCTCTAGACTGTCTTTG 780
QY 909 AGCCAGGCTGTAAACCTGACACTTTATATGTGTTTACACATGTAACTACATACACACA 968
Db 781 AGCCAGGCTGTAAACCTGACACTTTATATGTGTTTACACATGTAACTACATACACACA 840
QY 969 TGGCGCTCAGCAGCATGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1028
Db 841 TGGCGCTCAGCAGCATGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
QY 1029 CCCTTCTCAGGCTGTGTGATCTCTCTAGGGGATGGGGAGAGCCCTGGCTGCAGGCA 1088
```

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC007YC08"
/tissue_type="Neuroblastoma Cot 25-normalized"
/plasmid="pCMVSPORT_6"
```

Db 901 CTTCTCAGGCTGCTGATCTTCTTAGGAGTGGGGAAGCCCTGCTGAGGCA 960
 QY 1089 GCCTTCAGGCAATATGAGATAGAGGCCACCGGCTGCGAGTGAGAGGTGG 1144
 Db 961 GCCTTCAGGCAATATGAGATAGAGGCCACCGGCTGCGAGTGAGAGGTGG 1016

RESULT 4
 AL561088 1021 bp mRNA linear EST 02-APR-2004
 LOCUS AL561088 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
 DEFINITION Homo sapiens cDNA clone CS0DL001YJ06 5-PRIME, mRNA sequence.
 ACCESSION AL561088
 VERSION AL561088.3 GI:46186449
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1021)
 Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 15, 2001 this sequence version replaced gi:31285217.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
 Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 8673.f

For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?g=CS0DL001DE03QPI&c=8673.f>.

FEATURES

source

1..1021
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DL001YJ06"
 /cell_type="B CELLS (RAMOS CELL LINE)"
 /cell_line="RAMOS CELL LINE"
 /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 37.1%; Score 965.4; DB 1; Length 1021;
 Best Local Similarity 98.8%; Pred. No. 5.4e-201;
 Matches 977; Conservative 5; Mismatches 6; Indels 1; Gaps 1;
 QY 113 GCTACTAGATGGCGATGAGTTTCGAGTGGCGGAGTATCGGTTCCACCCCTCTTT 172
 Db 29 GCTACTAGATGGCGATGAGTTTCGAGTGGCGGAGTATCGGTTCCACCCCTCTTT 88
 QY 173 ACGTTTACACCGAATGTGACACTCGGCAGAGCAGCTGGCGCTGGTGTCTGCTGTC 232
 Db 89 ACGTTTACACCGAATGTGACACTCGGCAGAGCAGCTGGCGCTGGTGTCTGCTGTC 148
 QY 233 CTGCTCTTCTGCGGCTGCACAAACAGTCCAGCATGACGGTGTGATGGAAGCTCAGGAGC 292
 Db 149 CTGCTCTTCTGCGGCTGCACAAACAGTCCAGCATGACGGTGTGATGGAAGCTCAGGAGC 208
 QY 293 CGGCTCTTCAACAGCTCAAGCTACAGGAAAGCTTCTGTGGAGTCGATCCAGATTGTA 352
 Db 209 CGGCTCTTCAACAGCTCAAGCTACAGGAAAGCTTCTGTGGAGTCGATCCAGATTGTA 268

QY 353 TTAGAGGAACCTGAGAGAGAAAGGAAACCTCGAGTGGTTGGATAGAGCAAGTCCAGCTTC 412
 Db 269 TTAGAGGAACCTGAGAGAGAAAGGAAACCTCGAGTGGTTGGATAGAGCAAGTCCAGCTTC 328
 QY 413 CTGATCATGTGCGGAGGAGCCAGAGAAATGGGGAAACTCATCTATCAGTGGGTTTCCAGG 472
 Db 329 CTGATCATGTGCGGAGGAGCCAGAGAAATGGGGAAACTCATCTATCAGTGGGTTTCCAGG 388
 QY 473 AGTGGCCAGAAACAACCTCGTCTTTACCTGTATGAATGACTTAATGGGGAAGACACAG 532
 Db 389 AGTGGCCAGAAACAACCTCGTCTTTACCTGTATGAATGACTTAATGGGGAAGACACAG 448
 QY 533 GATGAGAGATTCCACGGGCTGATGAAGCACTCTACTGCGGGCTCTCAGGCCCTACAG 592
 Db 449 GATGAGAGATTCCACGGGCTGATGAAGCACTCTACTGCGGGCTCTCAGGCCCTACAG 508
 QY 593 CAGGAGCAAAAGGCGAGATCATCATGTCTAGCGATGCGCCGAGGCTCAAGTCTTCTTAG 652
 Db 509 CAGGAGCAAAAGGCGAGATCATCATGTCTAGCGATGCGCCGAGGCTCAAGTCTTCTTAG 568
 QY 653 CAGGAGCTGTCTCCCTTTACTTCTTACCTCCCACTTCCAGGGCTTTCAAAAGGAGAC 712
 Db 569 CAGGAGCTGTCTCCCTTTACTTCTTACCTCCCACTTCCAGGGCTTTCAAAAGGAGAC 628
 QY 713 AGACCCAGTGTCCCCCAAGACTGGATCTGTGACTCCACACAGACTCAAAAGGACTCCAGT 772
 Db 629 AGACCCAGTGTCCCCCAAGACTGGATCTGTGACTCCACACAGACTCAAAAGGACTCCAGT 688
 QY 773 CCTGAAGGCTGGGACCTGGGAGTGGGTTTCTCACACCCCATATGTCTCTCCCTGGATAG 832
 Db 689 CCTGAAGGCTGGGACCTGGGAGTGGGTTTCTCACACCCCATATGTCTCTCCCTGGATAG 748
 QY 833 GGTGAGGCTGAAGCACACGAGGAGAAATATGTGTTCTTCTCGCCCTACCTCTTCCCA 892
 Db 749 GGTGAGGCTGAAGCACACGAGGAGAAATATGTGTTCTTCTCGCCCTACCTCTTCCCA 808
 QY 893 TCCTAGACTGTCTTGAGCAGGCTGTGTAACCTGACACTTTATATGTGTTCACACATG 952
 Db 809 TCCTAGACTGTCTTGAGCAGGCTGTGTAACCTGACACTTTATATGTGTTCACACATG 868
 QY 953 TAAGTACATACACATGCGCTCGAGCAGATGTTCTGTCTCTCTCTCCACCCCT 1012
 Db 869 TAAGTACATACACATGCGCTCGAGCAGATGTTCTGTCTCTCTCTCTCCACCCCT 928
 QY 1013 TTAGTCTGTGTGCT 1072
 Db 929 TTAGTCTGTGTGCT 988
 QY 1073 GCCCTGGCTGAGGAGGAGGCTTCCAGGCAA 1101
 Db 989 G-CCTGGSTGCGAGSAGCYTTCAGGGA 1016

RESULT 5

BX357325/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BX357325 1079 bp mRNA linear EST 23-APR-2004
 BX357325 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 clone CS0DI024YM13 3-PRIME, mRNA sequence.
 BX357325
 BX357325.2 GI:46551167
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1079)
 Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 5, 2003 this sequence version replaced gi:30380160.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE

Query Match	35.4%	Score 923;	DB 5;	Length 1039;
Best Local Similarity	95.7%	Pred. NO.	1.1e-191;	
Matches 991; Conservative	0;	Mismatches 37;	Indels 7;	Gaps 4;

Qy	101	GGCAGCAGCTGGGGCTACTACGATGGCGAGTAGATTTCCGATGTGGCGCGTGGCAGTATCGCTTC	161
Db	5	GGGTGGGCGTGGGGCTACTACGATGGCGAGTAGATTTCCGATGTGGCGCGTGGCAGTATCGCTTC	64
Qy	161	CCACCCCTTTCTTTACGTTTACAAACCAGAACTGTGGACACTCTGGCAGAAAGCAGCTGGCGCCCTGG	220
Db	65	CCACCCCTTTCTTTACGTTTACAAACCAGAACTGTGGACACTCTGGCAGAAAGCAGCTGGCGCCCTGG	124
Qy	221	TGCTCGCTGGTCTCTGTCTCTTGGCGCCTGCAACAAAGTCCAGCATGACGGTATGATGAA	280
Db	125	TGCTCGCTGGTCTCTGTCTCTTGGCGCCTGCAACAAAGTCCAGCATGACGGTATGATGAA	184
Qy	281	GCTCAGGAGAGCCGCTCTTCAACAAACGTCGAAGCTACAGCGAAAGCTTCCTGTGTGAGTGC	340
Db	185	GCTCAGGAGAGCCGCTCTTCAACAAACGTCGAAGCTACAGCGAAAGCTTCCTGTGTGAGTGC	244
Qy	341	ATCCAGATTTGTAATTAGAGGAACCTGAGGAAGAAAGGAACTTCGAGTGGTTCGATAAGAGC	400
Db	245	ATCCAGATTTGTAATTAGAGGAACCTGAGGAAGAAAGGAACTTCGAGTGGTTCGATAAGAGC	304
Qy	401	AAGTCCAGCTTCTCTGATCATGTGGCGGAGCCAGAGAATGGGGAAACTCATCTATCAG	460
Db	305	AAGTCCAGCTTCTCTGATCATGTGGCGGAGCCAGAGAATGGGGAAACTCATCTATCAG	364
Qy	461	TGGGTTTCCAGGAGTGGCCAGAAACATCCGCTCTTTACCTGTATGAACCTGACTAATGGG	520
Db	365	TGGGTTTCCAGGAGTGGCCAGAAACATCCGCTCTTTACCTGTATGAACCTGACTAATGGG	424
Qy	521	GAAGACACAGAGGATGAGGAGTCCACGGGCTGGATGAAGCCACTCTACTCGGGGCTCTG	580
Db	425	GAAGACACAGAGGATGAGGAGTCCACGGGCTGGATGAAGCCACTCTACTCGGGGCTCTG	484
Qy	581	CAGGCCCTTACAGCAGGAGCACAAGGCCAGATCATCACTGTCAAGCATGGCCGAGGCGTC	640
Db	485	CAGGCCCTTACAGCAGGAGCACAAGGCCAGATCATCACTGTCAAGCATGGCCGAGGCGTC	544
Qy	641	AAGTTCTTTAGCAGGAGACCTGTCTCCCTTTACTTTTACCTCCCACTTTTCAGGGGCTT	700
Db	545	AAGTTCTTTAGCAGGAGACCTGTCTCCCTTTACTTTTACCTCCCACTTTTCAGGGGCTT	604
Qy	701	TCAAAAGGAGACAGACCCAGTGTCCCCAAAGACTGGATCTGTGACTCCACAGACTCAA	760
Db	605	TCAAAAGGAGACAGACCCAGTGTCCCCAAAGACTGGATCTGTGACTCCACAGACTCAA	664
Qy	761	AAGACTCCAGTCTCTGAAGGCTGGGACCTGGGGATGGGTTTCTCAACCCCATATGTCTG	820
Db	665	AAGACTCCAGTCTCTGAAGGCTGGGACCTGGGGATGGGTTTCTCAACCCCATATGTCTG	724
Qy	821	TCCCTTGGATAGGTTGAGGCTGAAGCACCAGGAGAAATATGTCTTCTCTCGCCCTA	880
Db	725	TCCCTTGGATAGGTTGAGGCTGAAGCACCAGGAGAAATATGTCTTCTCTCGCCCTA	784
Qy	881	CCTCCTTTCCCATCCTAGACTGTCTCTGAGCCAGGGCTGTAAACCTGACAC	938
Db	785	CCTCCTTTCCCATCCTAGACTGTCTCTGAGCCAGGGCTGTAAACCTGACACTTTATAT	844
Qy	939	TGTTGTTACACATGTAAGTACATACACATGCGCCTGCAGCATGCTTCTGTCTCTC	998
Db	845	GTTGTTACACATGTAAGTACATACACATGCGCCTGCAGCATGCTTCTGTGTTTCTC	904
Qy	999	CTCCTCCGACCCCTTTAGCTGCTGTGTGCTCTCCTTCTCAGGCTGGTGTGTGATCCTTCT	1058
Db	905	CTCCT-CCACCCCTTTAGCTGCTGTGTGCTCTCCTTCTCAGGCTGGTGTGTGATCCTTCTC	963
Qy	1059	A---GGGAGTGGGGAGCCCTGGCTGCAGG-CAGCCCTTCAGGCAATATGAAGATAGGA	1114
Db	964	AAGGGGGATGGGGGACCCCTGGCTTCAGGCCAGCCCTTCAGGAAATTTGAAATGAA	1023
Qy	1115	GGCCCCAGGGCCTGG	1129
Db	1024	GGCCCCGGGCCCTGG	1038

```

RESULT 7
AL582492/c      AL582492          971 bp      mRNA      linear      EST 07-APR-2004
LOCUS           AL582492 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION      Homo sapiens cDNA clone CS0DL001YJ06 3-PRIME, mRNA sequence.
ACCESSION       AL582492
VERSION         AL582492.3 GI:46260956
KEYWORDS        EST.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 971)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 16, 2001 this sequence version replaced gi:31320709.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Creteau, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 8673.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DL001DR03NP1&c=8673.f.
Location/Qualifiers
1..971
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL001YJ06"
/call_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match      35.1%; Score 914.8; DB 1; Length 971;
Best Local Similarity 97.5%; Pred.No. 7e-190;
Matches 949; Conservative 9; Mismatches 12; Indels 3; Gaps 3;
Qy 177 TACAACCGAATGTGGACACTCGGCAGAGCAGCTGCGCCGCTGGTGCTCCTCGTCTCTGT 236
Db |||:|||||||:|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
971 TAAACCGAATGT-GACACTCGGCAGAAGCRGCTGGCCGCTGGTGCTCGC-KGTCTCTGT 914
Qy 237 CCTTCTGCCGCTGCACAAACAGTCACAGCATGACGGTGATGGAGCTCACGAGAGCCCCG 296
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
913 CCTTCTGCCGCTGCACATACAGTCACAGCATGACGGTGATGGAGCTCACAGAGGCCCCG 854
Qy 297 TCATTCAACACGTCACAGCTACAGCGAAGCTTCTGTGGAGTCGATCCAGATTGTATTAG 356
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
853 TCATTCAACACGTCACAGCTACAGCGAAGCTTCTGTGGAGTCGATCCAGATTGTATTAG 794
Qy 357 AGGAACGTGAGAGAGAAAGGAACTCGAGTGGTTGGTAAGAGCAAGTCAGCTTCTCTGA 416
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
793 AGGSACTGTAGAGAGAAAGGAACTCGAGTGGTTGGTAAGAGCAAGTCAGCTTCTCTGW 734
Qy 417 TCATGTGCGGAGGCCAAGAGATGGGGAAACTCATCTATCAGTGGGTTTTCCAGAGAGTG 476
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
733 TCTTGTGGCGAGGCCAAGAGATGGGGGACATCATCTATCAGTGGGTTTCAGAGAGTG 674
Qy 477 GCCAGAACACTCGCTTTTACCCTGTATGACTACTATAATGGGGAGACACAGAGAGTG 536
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
673 GCCAGCAACTCGCTTTTACCCTGTATGTACTGACTAATGGGGBAGACACAGAGAGTG 614

```

FEATURES		Location/Qualifiers	
source		1..1005	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:5742295"	
		/tissue_type="medulla"	
		/lab_host="DH10B"	
		/clone_lib="NIH_MGC_119"	
		/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."	
ORIGIN		Query Match	
		Best Local Similarity 34.7%; Score 903; DB 4; Length 1005;	
		Matches 973; Conservative 0; Mismatches 21; Indels 8; Gaps 5;	
QY	114	CTACTACGATGCGGATGATGTTTCGAGTGGCCGCTGGGAGTATCGCTTCCACCCCTCTTTA	173
Db	1	CTACTACGATGCGGATGATGTTTCGAGTGGCCGCTGGGAGTATCGCTTCCACCCCTCTTTA	60
QY	174	CGTTACACCGAATGTGACACTCGGCAGAGACAGCTGCGTGGTATGGAAGCTCAGGAGAGCC	233
Db	61	CGTTACACCGAATGTGACACTCGGCAGAGACAGCTGCGCCGCTGGTCTCGTGGTCC	120
QY	234	TGTCCTTCTGCGCCCTGCACAAACAGTCAGCATGACGGTGTATGGAAGCTCAGGAGAGCC	293
Db	121	TGTCCTTCTGCGCCCTGCACAAACAGTCAGCATGACGGTGTATGGAAGCTCAGGAGAGCC	180
QY	294	CGCTCTTCAACAACTCAAGCTACAGCGAAAGCTCTCTGTGGAGTGCATCCAGATTGTAT	353
Db	181	CGCTCTTCAACAACTCAAGCTACAGCGAAAGCTCTCTGTGGAGTGCATCCAGATTGTAT	240
QY	354	TAGAGGACTGAGGAGAAAGGGAACCTCGAGTGTGGATAGAGCAAGTCCAGCTTCC	413
Db	241	TAGAGGAACTGAGGAGAAAGGGAACCTCGAGTGTGGATAGAGCAAGTCCAGCTTCC	300
QY	414	TGATCATGTGCGGAGGCGCAGAAAGTATGGGAAACTCATCTATCAGTGGGTTTCCAGCA	473
Db	301	TGATCATGTGCGGAGGCGCAGAAAGTATGGGAAACTCATCTATCAGTGGGTTTCCAGCA	360
QY	474	GTGGCCAGAACTCCCTCTTTTACCTGTATGAACCTGATGAGTAACTGAGTAACTGAGG	533
Db	361	GTGGCCAGAACTCCCTCTTTTACCTGTATGAACCTGATGAGTAACTGAGGAACTGAGG	420
QY	534	ATGAGGAGTTCACCGGGCTGGATGAGCCACTCTACTGCGGGCTCTGCGAGGCTTACAGC	593
Db	421	ATGAGGAGTTCACCGGGCTGGATGAGCCACTCTACTGCGGGCTCTGCGAGGCTTACAGC	480
QY	594	AGGAGCACAGGCCGAGATCATCTGTCAGCGATGGCGAGGCGTCAAGTCTTCTTACG	653
Db	481	AGGAGCACAGGCCGAGATCATCTGTCAGCGATGGCGAGGCGTCAAGTCTTCTTACG	540
QY	654	AGGAGCTGTCTCTCTTTTACCTTCTTACCTCCCACTTTCAGGGCTTTTCAAAAGGAGCA	713
Db	541	AGGAGCTGTCTCTCTTTTACCTTCTTACCTCCCACTTTCAGGGCTTTTCAAAAGGAGCA	600
QY	714	GACCCAGTGTCCCCCAAGACTGGATCTGTGATCCACGAGCTCAAGAGGACTCCAGTCC	773
Db	601	GACCCAGTGTCCCCCAAGACTGGATCTGTGATCCACGAGCTCAAGAGGACTCCAGTCC	660
QY	774	CTGAAGGCTGGGACCTGGGGATGGGTTTCTCACACCCCATATGTCTGTCTTGGATAGG	833
Db	661	CTGAAGGCTGGGACCTGGGGATGGGTTTCTCACACCCCATATGTCTGTCTTGGATAGG	720
QY	834	CTGAGGCTGAAGCACACGAGGAGAAATATGTGCTTTCTTCTGCGCCCTTACCTCTTCCCAT	893

```
Db      721 GTGAGCTGAAGCACCAGGAGAAATATGTCTTCTTCGCCCTACTCCTCTTCCCAT 780
Qy      894 CTTAGACTGTCTTCTGAGGCA-GGGTCTGTAAACCTGACACTTTATATGTGTTCACATG 952
Db      781 CTTAGACTGTCTTCTGAGCGAGGGTCTGTAAACCTGACACTTTATATGTGTTCACATG 840
Qy      953 TAAGTACATACACATCGGCTGAGCAGACATGCTTCTGTGTC-TCTCTCTCTCTCCACACCC 1011
Db      841 TAAGTACATACACATCGGCTGAGCAGACATGCTTTTGTGTTTCTCTCTTCTTCCCAACC 900
Qy      1012 TTTAGCTCTGTGTGCTCTCCCTT--CTCAGGCTGTGTGCTGATGCCCTCC--TAGGGGATG 1066
Db      901 TTTAGCTCTGTGTGCTCTCCCTTCTCAGGCTGTGTGCTGGAATCCTTCTTAAGGGGAGG 960
Qy      1067 GGGGAAGCCCTGCTGTCAGGAG-CCTTCAGGCAATATGAA 1107
Db      961 GGGAAACCCCTGGCTGGAAGCAGCCCTTCCAGGCAATTGAA 1002

RESULT 9
CA487570 982 bp mRNA linear EST 14-NOV-2002
LOCUS AGNCOURT 10809280 MAPcL Homo sapiens cdna clone IMAGE:6719062 5',
DEFINITION mRNA sequence.
ACCESSION CA487570
VERSION CA487570.1 GI:24947279
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 982)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: sgapbs-remail.nih.gov
Tissue Procurement: Kristi A. Egland, Ira Pastan
cdna Library Preparation: Invitrogen Corp
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAW14276 row: f column: 22
High quality sequence stop: 712.
FEATURES             Location/Qualifiers
     source           1..982
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6719062"
                     /cell_lines="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
                     hTERT-HME1, LNCap"
                     /lab_host="EMDH108"
                     /clone_lib="MAPcL"
                     /note="Vector: pCMV-SPORT6; Site 1: EcoRV; Site 2: Not I;
                     Subtracted with brain, liver, kidney and muscle.
                     Directionally cloned. Priming method: oligo-dT. Average
                     insert size: 1800 bp. Library amplification: 26,000 fold.
                     Kristi A. Egland, James J. Vincent, Robert Strausberg,
                     Bungkok Lee & Ira Pastan: Discovery of new breast
                     cancer genes encoding membrane and secreted proteins.
                     Manuscript submitted."
```

ORIGIN

```
Query Match      34.1%; Score 889; DB 6; Length 982;
Best Local Similarity 96.3%; Pred. No. 3.2e-184;
Matches 941; Conservative 0; Mismatches 29; Indels 7; Gaps 3;
Qy      133 TTTTCAGTGGCGTGGCAGTATCGCTTCCACCTCTTTAGCTTACACCAAGTGTGGA 192
```

```
Db      1 TTTTCAGTGGCGTGGCAGTATCGCTTCCACCTCTTTTACGTTTCAACAAGATGTGGA 60
Qy      193 CACTCGGCAGAACAGCTGGCGGCTGTGTCTCGTGTCTGTCTCTTCTGCGGCTGCA 252
Db      61 CACTCGGCAGAACAGCTGGCGGCTGTGTCTCGTGTCTGTCTCTTCTGCGGCTGCA 120
Qy      253 CAACAGCTCCAGCATGACGGTGTAGTGAAGCTTCAGGAGAGCCGCTCTTCAACAAGCTCAA 312
Db      121 CAACAGCTCCAGCATGACGGTGTAGTGAAGCTTCAGGAGAGCCGCTCTTCAACAAGCTCAA 180
Qy      313 GCTACAGCAAGCTTCTCTGTGGAGTCGATCCAGATTGTTATAGAGAACTGAGAAAGAA 372
Db      181 GCTACAGCAAGCTTCTCTGTGGAGTCGATCCAGATTGTTATAGAGAACTGAGAAAGAA 240
Qy      373 AGGGAACCTTCAGTGGTGGATAGAGCAAGTTCAGCTTCTCTGATCATGTGCGGAGGCC 432
Db      241 AGGGAACCTTCAGTGGTGGATAGAGCAAGTTCAGCTTCTCTGATCATGTGCGGAGGCC 300
Qy      433 AGAAGATGGGGGAAACTCATCTATCATGTGGTTCAGAGGATGGCCAGAACTCCGT 492
Db      301 AGAAGATGGGGGAAACTCATCTATCATGTGGTTCAGAGGATGGCCAGAACTCCGT 360
Qy      493 CTTTACCTGTATGAATGACTTAATGGGAAAGACACAGAGATGAGAGATTCACAGGGCT 552
Db      361 CTTTACCTGTATGAATGACTTAATGGGAAAGACACAGAGATGAGAGATTCACAGGGCT 420
Qy      553 GGATGAGCCACTTCTACTGCGGGCTCTCAGGCGCTCAGCAGGCTCAGCAGGAGCAAGCCGAGAT 612
Db      421 GGATGAGCCACTTCTACTGCGGGCTCTCAGGCGCTCAGCAGGAGCAAGCCGAGAT 480
Qy      613 CATCACTGTAGCGATGGCCGAGGCGTCAAGTTCCTTCTAGCAGGAGCACTGTCTCCCTTTA 672
Db      481 CATCACTGTAGCGATGGCCGAGGCGTCAAGTTCCTTCTAGCAGGAGCACTGTCTCCCTTTA 540
Qy      673 CTTCTTACCTCCCACTTTCAGGGCTTTCAAAAGAGAGACAGACCACTGTCTCCCAAG 732
Db      541 CTTCTTACCTCCCACTTTCAGGGCTTTCAAAAGAGAGACAGACCACTGTCTCCCAAG 600
Qy      733 ACTGGATCTGTGACTTCCACAGACTCAAAAGAGCTCCAGTCTTGAAGCTGGGACCTGGG 792
Db      601 ACTGGATCTGTGACTTCCACAGACTCAAAAGAGCTCCAGTCTTGAAGCTGGGACCTGGG 660
Qy      793 GATGGGTTTCTCACACCCCATATGTCTCTCTTGGATAGGCTGAGGCTGGAAGCACCAGG 852
Db      661 GATGGGTTTCTCACACCCCATATGTCTCTCTTGGATAGGCTGAGGCTGGAAGCACCAGG 720
Qy      853 GAGAAATATGTGTCTTCTGCGGCTTACCTCTCTTCCATCTCTAGACTGTCTTGAGCC 912
Db      721 GAGAAATATGTGTCTTCTGCGGCTTACCTCTCTTCCATCTCTAGACTGTCTTGAGCC 780
Qy      913 AGGGTCTGTAAACCTGACA-CTTTATATGTGTTCACACATGTAACTATACACACATGC 971
Db      781 AGGGTCTGTAAACCTGACACCTTTATATGTGTTCACACATGTAACTATACACACATGC 840
Qy      972 GCCTGCAGCACATGCTTCTGTCTCTCTCTCTCTCTCTCTCT-CCACNNCTTTAGCTGTGTGCTCCC 1031
Db      841 GCCTGCAGCACATGCTTCTGTCTCTCTCTCTCTCTCTCTCT-CCACNNCTTTAGCTGTGTGCTCCC 899
Qy      1032 TTCTCAGGCTGGTGTGATCTTCC-----TAGGGGATGGGGGAGCCCTGGCTGCAGG 1086
Db      900 TCTTCAGGCTGGGCTGATCTCTCTTTCCCAAGGGGATGGGGGAGCCCTGGGCTGCAGG 959
Qy      1087 CAGCTTCCAGGCAATA 1103
Db      960 CAGCCTTTCAAGCAAAA 976

RESULT 10
BX383031/c
LOCUS BX383031 Homo sapiens T CELLS (JURKAT CELL LINE) cOT 10-NORMALIZED
DEFINITION Homo sapiens cdna clone CSDDJ004YF19 3-PRIME, mRNA sequence.
ACCESSION BX383031
```


Db	122	GGCGCTGCACAAACAGTCCAGATGACGGTGTGAGAGCTCAGGAGACCCGCTCTTCA	181
Qy	303	ACAACGTCAAGCTACACGAAAGCTCTCTGTGGAGTTCGATCCAGATGTTATTAGAGGAC	362
Db	182	ACAACGTCAAGCTACACGAAAGCTCTCTGTGGAGTTCGATCCAGATGTTATTAGAGGAC	241
Qy	363	TCAGGAGAAAGGACCTCGAGTGGTTCGATAAGAGCAAGTCCAGCTTCTGATCATGT	422
Db	242	TCAGGAGAAAGGACCTCGAGTGGTTCGATAAGAGCAAGTCTAGCTTCTGATCATGT	301
Qy	423	GGCGAGGCCAGAGATGGGGAACTCATCTATCAGTGGTTCGAGAGTGGCCACA	482
Db	302	GGCGAGGCCAGAGATGGGGAACTCATCTATCAGTGGTTCGAGAGTGGCCACA	361
Qy	483	ACAACCTCCGTCTTTACCTGTATGAACCTGAATAGGGGAAGACACAGAGGATGAGGAT	542
Db	362	ACAACCTCCGTCTTTACCTGTATGAACCTGAATAGGGGAAGACACAGAGGATGAGGAT	421
Qy	543	TCACGGGTGATGAGGACCTACTCTCGGGCTCTGAGGCTCTACAGCAGGAGCACAC	602
Db	422	TCATGGGTGATGAGGACCTACTCTCGGGCTCTGAGGCTCTACAGCAGGAGCACAC	481
Qy	603	AGCGCAGATCATCTGTCAGCGATGGCGGAGGCTCAAGTCTTCTAGCAGGACCTG	662
Db	482	AGCGCAGATCATCTGTCAGCGATGGCGGAGGCTCAAGTCTTCTAGCAGGACCTG	541
Qy	663	TCTCCCTTTACTTTTACCTCCACCTTTCAGGGCTTTCAAAAGGAGACAGACCCAGTG	722
Db	542	TCTCCCTTTACTTTTACCTCCACCTTTCAGGGCTTTCAAAAGGAGACAGACCCAGTG	601
Qy	723	TCCCCAAAGACTGATCTGTGACTCCACAGACTCAAAAGGACTCCAGTCTGAAGGCT	782
Db	602	TCCCCAAAGACTGATCTGTGACTCCACAGACTCAAAAGGACTCCAGTCTGAAGGCT	661
Qy	783	GGGACTGGGATGGTTCCTCACCCCATATGTCCTGCTGGATAGGGTGGGCTG	842
Db	662	GGGACTGGGATGGTTCCTCACCCCATATGTCCTGCTGGATAGGGTGGGCTG	720
Qy	843	AAGCACCAGGAGAAAATATGTCTTCTCGCCCTACTCTCTTCCCATCTAGACTG	902
Db	721	AAGCACCAGGAGAAAATATGTCTTCTCGCCCTACTCTCTTCCCATCTAGACTG	779
Qy	903	TCCTTGAGCCAGGCTGTAAACCTGACACTTTATATGTTTCACACATGTAAGTACATA	962
Db	780	TCCTTGAGCCAGGCTGTAT-AACTGATCTTTATATGTTTCACACATGTAAGTACATA	838
Qy	963	CACATGCGCTGAGCAGCATGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1022
Db	839	CACATGCGCTGAGCAGCATGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	898
Qy	1023	TTGCT	1057
Db	899	TTGCT	933
RESULT 13.			
BM919274			
LOCUS			
DEFINITION			
5', mRNA Sequence.			
AGENCY			
BM919274			
VERSION			
BM919274.1			
KEYWORDS			
EST.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE			
1 (bases 1 to 1049)			
AUTHORS			
NIH-MGC http://mgi.nci.nih.gov/ .			
TITLE			
National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL			
Unpublished (1999)			
COMMENT			
Contact: Robert Strausberg, Ph.D.			
Email: cgabbs-r@mail.nih.gov			

Tissue Procurement: Life Technologies, Inc.			
cDNA Library Preparation: Life Technologies, Inc.			
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: Agencourt Bioscience Corporation			
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov			
Plate: LLAM12776 row: 9 column: 05			
High quality sequence stop: 650.			
FEATURES			
Location/Qualifiers			
1..1049			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:5748508"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_120"			
/note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH_MGC Library."			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
32.8%; Score 854.6; DB 5; Length 1049;			
97.8%; Pred. No. 1.2e-176; Indels 6; Gaps 5;			
Conservative 0; Mismatches 15;			
Qy	110	TGGGCTACTACGATGGCGATGAGTTTTCGAGTGGCGGTGGCAGTATCGCTTCCACCCCTTC	169
Db	1	TGGGCTACTACGATGGCGATGAGTTTTCGAGTGGCGGTGGCAGTATCGCTTCCACCCCTTC	60
Qy	170	TTTACGTTACACCGAATGTGGACACTCTGGCAGAGAGAGCTGGCGCCCTGGTGTCTCGCTG	229
Db	61	TTTACGTTACACCGAATGTGGACACTCTGGCAGAGAGAGCTGGCGCCCTGGTGTCTCGCTG	120
Qy	230	GTCCGTCTCTTCGCGCTGCACAAACAGTCCAGCATGACCGTGTGATGAAGCTCAGGAG	289
Db	121	GTCCGTCTCTTCGCGCTGCACAAACAGTCCAGCATGACCGTGTGATGAAGCTCAGGAG	180
Qy	290	AGCCCGCTCTTCAACAAAGCTCAAGCTACAGCGAAAGCTTCTCTGTGGAGTCGATCCAGATT	349
Db	181	AGCCCGCTCTTCAACAAAGCTCAAGCTACAGCGAAAGCTTCTCTGTGGAGTCGATCCAGATT	240
Qy	350	GTATTAGAGAACTCAGGAGAAAGGAACTCTCGAGTGGTGGATAAGAGCAAGTCCAGC	409
Db	241	GTATTAGAGAACTCAGGAGAAAGGAACTCTCGAGTGGTGGATAAGAGCAAGTCCAGC	300
Qy	410	TTCCGTGATCTGGCGGAGGCCAGAGAACTGGGGGAACTCATCTATCAGTGGGTTTCC	469
Db	301	TTCCGTGATCTGGCGGAGGCCAGAGAACTGGGGGAACTCATCTATCAGTGGGTTTCC	360
Qy	470	AGGAGTGGCCAGAACTCCGCTCTTTACCTGTATGAACCTGACTAATGGGGAAGACACA	529
Db	361	AGGAGTGGCCAGAACTCCGCTCTTTACCTGTATGAACCTGACTAATGGGGAAGACACA	420
Qy	530	GAGGATGAGGATTCACGGGCTGATGAAGCACTTACTCGGGGCTCTGAGAGCCCTA	589
Db	421	GAGGATGAGGATTCACGGGCTGATGAAGCACTTACTCGGGGCTCTGAGAGCCCTA	480
Qy	590	CAGCAGGAGCAAGCGCGAGATCATCTGTCCAGGATGGCGGCGCTCAAGTTCTTC	649
Db	481	CAGCAGGAGCAAGCGCGAGATCATCTGTCCAGGATGGCGGCGCTCAAGTTCTTC	540
Qy	650	TAGCAGGAGCACTGTCTCCCTTTACTTCTTACCTCCACCTTTCCAGGGCTTTCAAAGA	709
Db	541	TAGCAGGAGCACTGTCTCCCTTTACTTCTTACCTCCACCTTTCCAGGGCTTTCAAAGA	600
Qy	710	GACAGACCCAGTGTCCCCCAAGACTGGATCTGTGATCTCCACAGACTCAAAGGACTCC	769

```
|||||
601 GACAGACCCAGGTGTCCTCCCAAGAGCTGGATCTGTGACTCCACAGACTCAAAAGGACTCC 660
Db
770 AGTCTCGAAGCGTGGGACCTGGGGATGGGTTTCTCACACCCCATATGTGTGCTGCTTGGGA 829
Db
661 AGTCTCGAAGCGTGGGACCTGGGGATGGGTTTCTCACAC-CCCATATGTGTGCTGCTTGGGA 719
Qy
830 TAGGGTGAAGCTGAAGCACACAGGAGAGAAATATGTGTTCTTCGCGCCTACCTCCTTTC 889
Db
720 TAGGGTGAAGCTGAAGC-CCAGGGAGAGAAATATGTGTTCTTCGCGCCTACCTCCTTTC 778
Qy
890 CCATCTAGACTGTCTCTTGAGCCAGGGTCTGTAAACCTTGACACTTTATATGTGTTTCACAC 949
Db
779 CCTCTAGACTGTCTCTTGAGCCAGGCTGTGTAAACCTTGACACTTTATATGTGTTCCAC 838
Qy
950 ATGTAAGTACATACACACATGGG-CCTGCAGACATGCTTCT-CTCTCCTCCTCCTCCC 1006
Db
839 ATGTAAGAACATACACCATCGCCCTGCAGCACATGCTTCTTGTNTCCCTCCTCCTCCC 898
Qy
1007 ACCCCTTTAGCTGCT-GTTGCTCCTCTCTCAGGCTGGTG 1045
Db
899 ACCCCTTTAAGCTGGTGGCTCCTCTCTCGGGCTGGGG 938
```

```
RESULT 14
BG675348          925 bp      mRNA      linear      EST 01-MAY-2001
LOCUS             602621657F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4754991 5',
DEFINITION        mRNA sequence.
ACCESSION         BG675348
VERSION           BG675348.1 GI:13906744
KEYWORDS          EST.
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens
REFERENCE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE             NIH-MGC http://mgi.nci.nih.gov/
JOURNAL           National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT           Unpublished (1999)
                  Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-remail.nih.gov
                  Tissue Procurement: James Cleaver, M.D.
                  cDNA Library Preparation: Life Technologies, Inc.
                  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
                  Sequencing by: Incyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: LLAM10617 row: b column: 16
                  High quality sequence stop: 838.
FEATURES          Location/Qualifiers
                  1..925
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:4754991"
                    /lab_host="DH10B (T1 phage-resistant)"
                    /clone_lib="NCI_CGAP_Skn3"
                    /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
                    Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
                    Average insert size 1.5kb. Library constructed by Life
                    Technologies. Note: this is a NCI_CGAP Library."
```

FEATURES

source

ORIGIN

```
Query Match      32.4%; Score 844.8; DB 4; Length 925;
Best Local Similarity 98.6%; Pred. No. 1.6e-174;
Matches 915; Conservative 0; Mismatches 7; Indels 6; Gaps 6;

Qy 136 CGAGTGGCGGTGGCAGTATCGTTCACCCCTTTTACGTTACACCGAA-TGTGGACA 194
Db 1 CGAGTGGCGGTGGCAGTATCGTTCACCCCTTTTACGTTACACCGAAAGTGTGGACA 60
```

```
Qy 195 CTGGCAGAGCAGCTGGCCGCTGGTGTCTCGCTGGTCTCTGTCTCTCTGCGCGCTGCACA 254
Db
61 CTGGCAGAGCAGCTGGCCGCTGGTGTCTCGCTGGTCTCTGTCTCTCTGCGCGCTGCACA 120
Qy 255 AACAGTCCAGCATGACGGTGATGGAAGCTCAGGAGAGCCCGCTCTTTCAACAACGCTCAAGC 314
Db
121 AACAGTCCAGCATGACGGTGATGGAAGCTCAGGAGAGCCCGCTCTTTCAACAACGCTCAAGC 180
Qy 315 TACAGCGAAGCTTCCTGTGAGTCGATCCAGATGTATTAGAGAACTGAGGAAAGAAAG 374
Db
181 TACAGCGAAGCTTCCTGTGAGTCGATCCAGATGTATTAGAGAACTGAGGAAAGAAAG 240
Qy 375 GGAACCTCGAGTGGTTGATTAAGAGCAAGTCCAGCTTCCTGATCATGTGGCGGAGGCCAG 434
Db
241 GGAACCTCGAGTGGTTGATTAAGAGCAAGTCCAGCTTCCTGATCATGTGGCGGAGGCCAG 300
Qy 435 AAGAAATGGGGGAAAACCTCATCTATCATGTGGGTTTCCAGAGTGGCCAGAACAACTCCGCT 494
Db
301 AAGAAATGGGGGAAAACCTCATCTATCATGTGGGTTTCCAGAGTGGCCAGAACAACTCCGCT 360
Qy 495 TTACCTCTGTATGAACCTGACTTAATGGGGAAGACACAGAGATGAGGAGTTCCACGGGCTGG 554
Db
361 TTACCTCTGTATGAACCTGACTTAATGGGGAAGACACAGAGATGAGGAGTTCCACGGGCTGG 420
Qy 555 ATGAAGCCACTCTACTGCGGCTCTGCAGGCCCTACACAGAGACACAAGGCCAGAGATCA 614
Db
421 ATGAAGCCACTCTACTGCGGCTCTGCAGGCCCTACACAGAGACACAAGGCCAGAGATCA 480
Qy 615 TCACGTGCAGCGATGGCGAGGCGTCAAGTTCTTCTAGCAGGAGACCTGTCTCCCTTTACT 674
Db
481 TCACGTGCAGCGATGGCGAGGCGTCAAGTTCTTCTAGCAGGAGACCTGTCTCCCTTTACT 540
Qy 675 TCTTACCTCCACCTTTCAGGGCTTTCAAAGGAGACAGACCCAGTGTCCCCCAAGAC 734
Db
541 TCTTACCTCCACCTTTCAGGGCTTTCAAAGGAGACAGACCCAGTGTCCCCCAAGAC 600
Qy 735 TGGATCTGTGACTCCACCAGACTCAAAGAGACTCCAGTCTCGAAGGCTGGGACCTGGGGA 794
Db
601 TGGATCTGTGACTCCACCAGACTCAAAGAGACTCCAGTCTCGAAGGCT-GGACCTGGGGA 659
Qy 795 TGGGTTTCTCACACCCCATATGTCTGTCTTGTGATAGGGTGAAGCTGAAGCACACAGGA 854
Db
660 TGGGTTTCTCACACCCCATATATGTCTGTCTTGGATAGGGTGAAGCTGAAGCACACAGGA 719
Qy 855 GAAATATGTCTTCTCGCCCTACCTCC-TTTCCCATCTAGACTGTCTTTCAGGACA 913
Db
720 GAAATATGTCTTCTCGCCCTACCTCTCTTTTCCCATCTAGACTGTCTTTCAGGACA 779
Qy 914 GGGTCTGTAAACCTGACACTTTATATGTGTTTCACACAT-GTAAAGTACATACACATGCG 972
Db
780 GGGTCTGTAAACCTGACACTTTATATGTGTTTCACACATGTTAAGTACATACACATGCG 839
Qy 973 CTTGAGCACATGCTTGTCTCTCTCTCTCTCCACCCCTTTAGTGTGTGTGCTCCTCCT 1032
Db
840 CTTG-AGCACATGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 897
Qy 1033 TCTCAGGCTGGTGGTGGATCTCTCTAG 1060
Db
898 TCTCAGGATGGTGGTGGATCTCTCAAG 925

RESULT 15
BM555365
LOCUS             AGENCOURT_6541180 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5550385
DEFINITION        5', mRNA sequence.
ACCESSION         BM555365
VERSION           BM555365.1 GI:18795758
KEYWORDS          EST.
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

REFERENCE 1 (bases 1 to 990)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA library Preparation: Life Technologies, Inc.
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLMW2262 row: h column: 02
High quality sequence stop: 685.

FEATURES
source
1..990
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="IMAGE:550385"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="PH10B (phage-resistant)"
/clone_lib="NIH MGC 88"
/note="Organ: small intestine; Vector: pCMV-SPORT6; Site 1: Noti; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.767 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 32.4%; Score 843.8; DB 4; Length 990;
Best Local Similarity 98.5%; Pred. No. 2.7e-174;
Matches 871; Conservative 0; Mismatches 10; Indels 3; Gaps 2;
QY 153 ATCGCTTCCACCCCTCTTTAGCTTACACCGAATGTGGACACTCGGCAGAGACGCTGG 212
DB 1 ATCGCTTCCACCCCTCTTTAGCTTACACCGAATGTGGACACTCGGCAGAGACGCTGG 60
QY 213 CGCCTGTGTCTCGCTGTCTCTTCTGCGGCTGCACAAAGTCCAGCATGAAGG 272
DB 61 CGCCTGTGTCTCGCTGTCTCTTCTGCGGCTGCACAAAGTCCAGCATGAAGG 120
QY 273 TGATGGAAGCTCAGGAGAGCCGCTTTCAACAAAGTCAAGTACAGGAAAGCTTCCTG 332
DB 121 TGATGGAAGCTCAGGAGAGCCGCTTTCAACAAAGTCAAGTACAGGAAAGCTTCCTG 180
QY 333 TGGAGTCGATCCAGATTGTATTAGGAGACTGAGGAGAAAGGAACTTCGAGTGGTGG 392
DB 181 TGGAGTCGATCCAGATTGTATTAGGAGAACTGAGGAGAAAGGAACTTCGAGTGGTGG 240
QY 393 ATAAGAGCAAGTCCAGCTTCTGATCATGTGCGGAGGCGCAGAGAAATGGGGAACTCA 452
DB 241 ATAAGAGCAAGTCCAGCTTCTGATCATGTGCGGAGGCGCAGAGAAATGGGGAACTCA 300
QY 453 TCTATCAGTGGGTTTCCAGGAGTGGCCAGAACTCCGTCTTTTACCCCTGTATGAAGTGA 512
DB 301 TCTATCAGTGGGTTTCCAGGAGTGGCCAGAACTCCGTCTTTTACCCCTGTATGAAGTGA 360
QY 513 CTAAATGGGAGACACAGAGGATGAGGAGTTCCAGGGCTGGATGAGCCACTTACTGC 572
DB 361 CTAAATGGGAGACACAGAGGATGAGGAGTTCCAGGGCTGGATGAGCCACTTACTGC 420
QY 573 GGGCTCTGCAGGCCCTTACAGAGGAGCAAGCGGAGATCATCTGTCCAGGATGGCC 632
DB 421 GGGCTCTGCAGGCCCTTACAGAGGAGCAAGCGGAGATCATCTCCGTCCAGGATGGCC 480
QY 633 GAGGGCTCAAGTTCTTTAGCAGGAGACCTGTCTCCCTTTTACTTTTACCTCCACCTTTC 692
DB 481 GAGGGCTCAAGTTCTTTAGCAGGAGACCTGTCTCCCTTTTACTTTTACCTCCACCTTTC 540
QY 693 CAGGGCTTCAAAGAGACAGACCCAGTGTCCCCCAAGAGCTGGATCTGTGACTCCACC 752
DB 1 CAGGGCTTCAAAGAGACAGACCCAGTGTCCCCCAAGAGCTGGATCTGTGACTCCACC 600
QY 753 AGACTCAAAGAGACTCCAGTCTCTGAAGGCTGGGACCTGGGGATGGGTTTCTCACAACCCCA 812
DB 601 AGACTCAAAGAGACTCCAGTCTCTGAAGGCTGGGACCTGGGGATGGGTTTCTCACAACCCCA 660
QY 813 TATGTCTGTCTCTTGGATAGGCTGAGGCTGGAAGCACCAGGGAGAAATATGTGCTTCTTC 872
DB 661 TATGTCTGTCTCTTGGATAGGCTGAGGCTGGAAGCACCAGGGAGAAATATGTGCTTCTTC 720
QY 873 TCGCCCTACCTCTCTTCCCATCTTAGACTGTCTTTGAGCCAGGGTCTGTAAACCTTGACAC 932
DB 721 TCGCCCTACCTCTCTTCCCATCTTAGACTGTCTTTGAGCCAGGGTCTGTAAACCTTGACAC 780
QY 933 TTTATATGTGTTTCAACATGTAAAGTACATACACAT-GCGCCTGCAGCACATGCTT--C 989
DB 781 TTTATATGTGTTTCAACATGTAAAGTACATACACATGCGCGCTGCAGCACATGCTTCT 840
QY 990 TGTCTCTCTCTCTCTCCGACCCCTTTAGCTGTGTTGCTTCCCTT 1033
DB 841 GTNNTCTCTCTCTCTCTCCACCCCTTTAGCTGTGTTGCTTCCCTT 884

Search completed: June 29, 2005, 05:14:46
JOB time : 5615.03 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 14:57:12 ; Search time 8104.93 Seconds
(without alignments)
16865.305 Million cell updates/sec

Title: US-08-731-499-6
Perfect score: 2821
Sequence: 1 ATCTTAAGACGACAGCCTG.....AGGGTTCAAGACCAGCCTG 2821

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2821	100.0	2821	6	AR070329
2	2821	100.0	2821	6	BD085730
3	2805.8	99.5	167078	9	HSJ139D8
4	1679.6	59.5	3008	9	HSGCAP13
5	1635.2	58.0	2288	9	HSW805690
6	956.6	33.9	2967	9	HSW807325
7	374	13.3	407	11	G30061
8	337	11.9	1506	6	CQ491451
9	337	11.9	1506	6	CQ497332
10	278.6	9.9	333201	2	AL390735
11	278	9.9	146530	2	AC023158
12	277.8	9.8	74754	2	CR759924
13	277.8	9.8	78827	9	CR788282
14	277.8	9.8	109067	9	CR388382
15	277.8	9.8	148281	2	CR759926
16	277.8	9.8	187607	9	AC010329
17	277.4	9.8	154957	9	AC019270
18	277.2	9.8	56403	9	HS4418064
19	277.2	9.8	164076	9	AC021059

20	276.8	9.8	172981	9	AP000851	AP000851 Homo sapi
21	276.8	9.8	185323	2	AC078970	AC078970 Homo sapi
22	276.4	9.8	105580	9	HS36J3	Z82975 Human DNA s
23	276	9.8	138411	9	HSJ71H19	AL121958 Human DNA
24	275.8	9.8	71092	9	AL590732	AL590732 Human DNA
25	275.8	9.8	83727	9	AL356120	AL356120 Human DNA
26	275.8	9.8	114701	9	AC006927	AC006927 Homo sapi
27	275.6	9.8	160396	9	AL592063	AL592063 Human DNA
28	275.6	9.8	187350	9	AC020713	AC020713 Homo sapi
29	275.2	9.8	173260	9	CNS078CU	AL442183 Human chr
30	275	9.7	38601	9	AC112646	AC112646 Homo sapi
31	275	9.7	45120	9	AC090156	AC090156 Homo sapi
32	275	9.7	114001	9	HSJ760G15	AL079340 Human DNA
33	275	9.7	161943	9	HS94M10	AL356460 Homo sapi
34	275	9.7	165799	2	AC100813	AC100813 Homo sapi
35	275	9.7	174373	9	AC132216	AC132216 Homo sapi
36	274.6	9.7	111783	9	AL592449	AL592449 Human DNA
37	274.6	9.7	144314	2	AP000835	AP000835 Homo sapi
38	274.6	9.7	171732	9	AC022844	AC022844 Homo sapi
39	274.6	9.7	176560	2	AC021852	AC021852 Homo sapi
40	274.6	9.7	180019	9	AC013734	AC013734 Homo sapi
41	274.6	9.7	180547	9	AL139331	AL139331 Human DNA
42	274.6	9.7	186152	2	AC130358	AC130358 Homo sapi
43	274.6	9.7	192070	9	AC090616	AC090616 Homo sapi
44	274.2	9.7	107685	9	BX679664	BX679664 Human DNA
45	274.2	9.7	129819	9	AC104049	AC104049 Homo sapi

ALIGNMENTS

RESULT 1
AR070329
LOCUS AR070329 2821 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 6 from patent US 5892010.
ACCESSION AR070329
VERSION AR070329.1 GI:7221217
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2821)
AUTHORS Gray, J., Collins, C., Hwang, S.-i., Godfrey, T., Kowbel, D. and Rommens, J.
TITLE Genes from the 20Q13 amplicon and their uses
JOURNAL Patent: US 5892010-A 6 06-APR-1999;
FEATURES
source Location/Qualifiers
1..2821
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN	Query Match	100.0%;	Score 2821;	DB 6;	Length 2821;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 2821;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Qy	1	ATCCTAAGACGACAGCCTGGGAAGCTGCTGGGGAAGTGGTCTGAGGATGTGGGT	60		
Db	1	ATCCTAAGACGACAGCCTGGGAAGCTGCTGGGGAAGTGGTCTGAGGATGTGGGT	60		
Qy	61	CACCTGGGTGAAGTGGAGCTTTTCAGGGTCTCCCGTCAATGACGTAGTCTTTTGGC	120		
Db	61	CACCTGGGTGAAGTGGAGCTTTTCAGGGTCTCCCGTCAATGACGTAGTCTTTTGGC	120		
Qy	121	AGGGAATTACCACTGAAGAGCTCCGCGGAGAGCTACAACTGACGAGCCAGC	180		
Db	121	AGGGAATTACCACTGAAGAGCTCCGCGGAGAGCTACAACTGACGAGCCAGC	180		
Qy	181	TGCTCACACCCGAGGAGTCTGGGACAGATCTTCTCTCTGTTGGATGAGATG	240		
Db	181	TGCTCACACCCGAGGAGTCTGGGACAGATCTTCTCTCTGTTGGATGAGATG	240		
Qy	241	GTAAGAGGGGACAGATGGGGAGAGTGTGTCTTCCACTCTGTCATCATCGCCACTTTCTGGCC	300		

Db 241 GTAAAGGGCAGAGATGGGAGAGTGTGTCCACTCTGCAATCAATCGCCACTTTCTGGCC 300
QY 301 GCACGCTCTTGGGCAAGGCCCTCACCCTTCCAAACCTCGGGTCTCATCTGTGAGAAAGC 360
Db 301 GCACGCTCTTGGGCAAGGCCCTCACCCTTCCAAACCTCGGGTCTCATCTGTGAGAAAGC 360
QY 361 TGTGAGAAGATGTCTATGAACCTAAACAAAGGACTCATAGACACGTGTTTGTAGGAGTGAC 420
Db 361 TGTGAGAAGATGTCTATGAACCTAAACAAAGGACTCATAGACACGTGTTTGTAGGAGTGAC 420
QY 421 TAAAGTCTTACAGAGTGTCTGATGGAGGCCAGGCACGCAGAAATAGAAAGATAGGAAC 480
Db 421 TAAAGTCTTACAGAGTGTCTGATGGAGGCCAGGCACGCAGAAATAGAAAGATAGGAAC 480
QY 481 TTTGAGTCAAGCAGGAGTGATATATTGAGCTTCTCGTCTTAGTCTCAATTTCTCATC 540
Db 481 TTTGAGTCAAGCAGGAGTGATATATTGAGCTTCTCGTCTTAGTCTCAATTTCTCATC 540
QY 541 TGGAAATGGGGATAATAATAGTGTGTGAGAGGAATGAATAGGATAATGTGTTTAAAGAGC 600
Db 541 TGGAAATGGGGATAATAATAGTGTGTGAGAGGAATGAATAGGATAATGTGTTTAAAGAGC 600
QY 601 AGGCATAGGGTAGACCTCCATTACAGCTGCTTTGGGCTTTTCTCCCTGTAGCCCAAGGCC 660
Db 601 AGGCATAGGGTAGACCTCCATTACAGCTGCTTTGGGCTTTTCTCCCTGTAGCCCAAGGCC 660
QY 661 AGCCTCAGGGTATGTGGGAGAGAGCTGGCTTGGAAATACACACTTGAGCCCTCAGCTC 720
Db 661 AGCCTCAGGGTATGTGGGAGAGAGCTGGCTTGGAAATACACACTTGAGCCCTCAGCTC 720
QY 721 TCTCAGCTCCACCCAGCATTTCCGTGTGTACCATGCGCAAAAGTAAATCTCAATTCATCA 780
Db 721 TCTCAGCTCCACCCAGCATTTCCGTGTGTACCATGCGCAAAAGTAAATCTCAATTCATCA 780
QY 781 GCAAAGAAAGCCCTTAAAGGTGGCAGGAGACTCTCTGGAGATTTACAGACCTCGACAAGCC 840
Db 781 GCAAAGAAAGCCCTTAAAGGTGGCAGGAGACTCTCTGGAGATTTACAGACCTCGACAAGCC 840
QY 841 GCAAGCTTGAGGCTGAGACTGAGACTGAGGATAGTTGGCATTAAGAGTGTAGCGCATCTGGG 900
Db 841 GCAAGCTTGAGGCTGAGACTGAGGATAGTTGGCATTAAGAGTGTAGCGCATCTCTGGG 900
QY 901 AGCAGAGTCTCTCTCTGCCCCCAGACCCAGCTCTCCCTTCTTCTACATGACCACTC 960
Db 901 AGCAGAGTCTCTCTCTGCCCCCAGACCCAGCTCTCCCTTCTTCTACATGACCACTC 960
QY 961 TCCTCCCCCTTGTCTCAGGCCAGCTGTCTCTGAACGAGTGTGTGAAGGTGCCCCGTCGGGA 1020
Db 961 TCCTCCCCCTTGTCTCAGGCCAGCTGTCTCTGAACGAGTGTGTGAAGGTGCCCCGTCGGGA 1020
QY 1021 CAAAGTGGGTGATGAAGATGCTGCAGATGGACATGAATCCACAGCAGCTGGCTCGCTCAGCA 1080
Db 1021 CAAAGTGGGTGATGAAGATGCTGCAGATGGACATGAATCCACAGCAGCTGGCTCGCTCAGCA 1080
QY 1081 GAGACGGAAAGTGCCATGTTCTGAGGAGTCTGGGGCCCTTCCACGACTCCAGGCTCAC 1140
Db 1081 GAGACGGAAAGTGCCATGTTCTGAGGAGTCTGGGGCCCTTCCACGACTCCAGGCTCAC 1140
QY 1141 CAGGTTTCCAGGGTAGTAGGAGGTCCTCGCTCAGCTGTGCTCATGCCCCACTCTTCCCC 1200
Db 1141 CAGGTTTCCAGGGTAGTAGGAGGTCCTCGCTCAGCTGTGCTCATGCCCCACTCTTCCCC 1200
QY 1201 TGGTGTGACTCTTCGGCACCCCTGTGCAGGGCTGAGTGGGGATGGGAGGGCTGCTG 1260
Db 1201 TGGTGTGACTCTTCGGCACCCCTGTGCAGGGCTGAGTGGGGATGGGAGGGCTGCTG 1260
QY 1261 GGTTTGAAGTGCCCAACAGGGCATAGTCCATTTTGGAGAGTCCCTCGGATGGTGAAGGG 1320
Db 1261 GGTTTGAAGTGCCCAACAGGGCATAGTCCATTTTGGAGAGTCCCTCGGATGGTGAAGGG 1320
QY 1321 AATTCAAGTACTTTTCTGTTTCAGCCGCTCTCGGAGGACTGTGCTTGGCTGGGTGGTT 1380

Db 1321 AATTCAAGTACTTTTCTGTTTCAGCCGCTCTCGGAGGACTGTGCTTGGCTGGGTGGTT 1380
QY 1381 GTGGGCTCCCAACAGTCTTCTCGGTGTTCTAGTGTGGAAGCAAGACCAACTGAGGGGTGA 1440
Db 1381 GTGGGCTCCCAACAGTCTTCTCGGTGTTCTAGTGTGGAAGCAAGACCAACTGAGGGGTGA 1440
QY 1441 GGGTCCCAACAGACCAAAATCAGAAATAGAAACACAAAGACTGTTAGGAGCGAGGGTGGGA 1500
Db 1441 GGGTCCCAACAGACCAAAATCAGAAATAGAAACACAAAGACTGTTAGGAGCGAGGGTGGGA 1500
QY 1501 GGGTGTGAGACTGAGAAAAAGGAGGAGTGTGCGGGGCAACGGTGGCTCAGCCTGTAATC 1560
Db 1501 GGGTGTGAGACTGAGAAAAAGGAGGAGTGTGCGGGGCAACGGTGGCTCAGCCTGTAATC 1560
QY 1561 CCAGCACTTTGGGAGGCCGAGCGGCGGAGATCAGAGGTGAGGAGATCAGAGCAACCATCTG 1620
Db 1561 CCAGCACTTTGGGAGGCCGAGCGGCGGAGATCAGAGGTGAGGAGATCAGAGCAACCATCTG 1620
QY 1621 GCTAACACGGGTGAAACCCCGTCTCTACTAAAAATACAAAAATCAGCCCGGTGAGGTG 1680
Db 1621 GCTAACACGGGTGAAACCCCGTCTCTACTAAAAATACAAAAATCAGCCCGGTGAGGTG 1680
QY 1681 GCGGGCGCTGTAGTCCAGCTACTCAGAGGCTGAGGCAAGAGAAATGGCGTGAACCCCA 1740
Db 1681 GCGGGCGCTGTAGTCCAGCTACTCAGAGGCTGAGGCAAGAGAAATGGCGTGAACCCCA 1740
QY 1741 GGGGGCGAGCCTACAGTGAGCGGAGATTTGCCCACTGCATCTCCAGCTTGGACGACAGTG 1800
Db 1741 GGGGGCGAGCCTACAGTGAGCGGAGATTTGCCCACTGCATCTCCAGCTTGGACGACAGTG 1800
QY 1801 AGACTCCGCTCTCAAAAAAAGAAAGAAAGAAAGGAGGAGGTTTTGGGGGGCAG 1860
Db 1801 AGACTCCGCTCTCAAAAAAAGAAAGAAAGAAAGGAGGAGGTTTTGGGGGGCAG 1860
QY 1861 GGGGAGCAATTAATTTCTATAACTTCCGGGATGCTGAGGGGCTTTCATGGGAGGACCCCTG 1920
Db 1861 GGGGAGCAATTAATTTCTATAACTTCCGGGATGCTGAGGGGCTTTCATGGGAGGACCCCTG 1920
QY 1921 GCCTCTCTCTCCCAAGGCATCTCACCAGTGTGTCAACAGGAAATGGCAGCAATA 1980
Db 1921 GCCTCTCTCTCCCAAGGCATCTCACCAGTGTGTCAACAGGAAATGGCAGCAATA 1980
QY 1981 CGCTGACGGCTGTGCTCTTTCTGCTTTGAAAGGCTCAGCTGTACTTAAAGGACTGT 2040
Db 1981 CGCTGACGGCTGTGCTCTTTCTGCTTTGAAAGGCTCAGCTGTACTTAAAGGACTGT 2040
QY 2041 CAGCTCTGCTGGGTGTGCTCTGGGACCCCTGCTGCAACCCACCACTCCCCCAACA 2100
Db 2041 CAGCTCTGCTGGGTGTGCTCTGGGACCCCTGCTGCAACCCACCACTCCCCCAACA 2100
QY 2101 TCCTCTCTTTCCATCCATATCCCGAGTATGGAACCTTCCAACTCCCGAGCATTAGCTG 2160
Db 2101 TCCTCTCTTTCCATCCATATCCCGAGTATGGAACCTTCCAACTCCCGAGCATTAGCTG 2160
QY 2161 AATGTTTCTTTTAAAGGATGGAGAAACTTCTGCTGTCTCTGCAAGAAATTTGGGGAC 2220
Db 2161 AATGTTTCTTTTAAAGGATGGAGAAACTTCTGCTGTCTCTGCAAGAAATTTGGGGAC 2220
QY 2221 TGTGACTGAGGATGTGGGCTGGGCTTGGCTTCTAACTGCTGTGACCAAGACAGCA 2280
Db 2221 TGTGACTGAGGATGTGGGCTGGGCTTGGCTTCTAACTGCTGTGACCAAGACAGCA 2280
QY 2281 CTTCTCTCCCTTAACTTTGGTTATGTCTTGGCAGCAGTGGAGGTCGGACTAGGCGA 2340
Db 2281 CTTCTCTCCCTTAACTTTGGTTATGTCTTGGCAGCAGTGGAGGTCGGACTAGGCGA 2340
QY 2341 ACAGTTTTGGATTATGTGTTTTAGATGTGGAATATTTTTTGTATATAAACTCTTAT 2400
Db 2341 ACAGTTTTGGATTATGTGTTTTAGATGTGGAATATTTTTTGTATATAAACTCTTAT 2400
QY 2401 GTGTAAACCCCAATATAGAAACTAGATTAAGGGAGTCTCTCTGTTGAAGGGGAGCTG 2460
Db 2401 GTGTAAACCCCAATATAGAAACTAGATTAAGGGAGTCTCTCTGTTGAAGGGGAGCTG 2460

1141 CAGGTTTCCAGGGTAGTAGGAGGGTCCCTCGCTCAGCCTGCTCATGCCCCACTCTTCCCC 1200
1141 CAGGTTTCCAGGGTAGTAGGAGGGTCCCTCGCTCAGCCTGCTCATGCCCCACTCTTCCCC 1200
1201 TGGTGTGACCTCTCTGGCACCCTCTGTGACGGCTGAGTGGGGATGGGGAAGGGCTGCTG 1260
1201 TGGTGTGACCTCTCTGGCACCCTCTGTGACGGCTGAGTGGGGATGGGGAAGGGCTGCTG 1260
1261 GGTTTGAAGTGGCCAAAGGGCATAGTCATTTTGGAGGAGTCCCTGGGATGGTGAAGGG 1320
1261 GGTTTGAAGTGGCCAAAGGGCATAGTCATTTTGGAGGAGTCCCTGGGATGGTGAAGGG 1320
1321 AATTGAGTTACTTTTCTGTTTCAGCCGCTCTCTGGGAGGACTGTGCTTGGCTGGGTGGTT 1380
1321 AATTGAGTTACTTTTCTGTTTCAGCCGCTCTCTGGGAGGACTGTGCTTGGCTGGGTGGTT 1380
1381 GTGGGGCTCCACAGATTTCTGGGTGTTCTCAAGTTGGAAGCAAGAGCCAACTGAGGGGTGA 1440
1381 GTGGGGCTCCACAGATTTCTGGGTGTTCTCAAGTTGGAAGCAAGAGCCAACTGAGGGGTGA 1440
1441 GGGTCCCAACAGCCAAATCAGAAATGAGAAACAAAGACTGGTAGGAGGCAAGGGGTGGGA 1500
1441 GGGTCCCAACAGCCAAATCAGAAATGAGAAACAAAGACTGGTAGGAGGCAAGGGGTGGGA 1500
1501 GGGTGTTCAGACTGAAGAAAAAGGCAGAGTTGCGGGGCACGCTGGCTCACGCCCTGTAATC 1560
1501 GGGTGTTCAGACTGAAGAAAAAGGCAGAGTTGCGGGGCACGCTGGCTCACGCCCTGTAATC 1560
1561 CCAGCAGTTTGGAGCCCGAGCGGGCAGATCAGAGGTCAGGAGATCGAGACCATCCTG 1620
1561 CCAGCAGTTTGGAGCCCGAGCGGGCAGATCAGAGGTCAGGAGATCGAGACCATCCTG 1620
1621 GCTAACACGGGTGAAACCCCGTCTCTACTAAAAATACAAAAATCAGCCGGGTGAGGTG 1680
1621 GCTAACACGGGTGAAACCCCGTCTCTACTAAAAATACAAAAATCAGCCGGGTGAGGTG 1680
1681 GCGGGCGCTGTAGTCCAGCTACTCAGAGGCTGAGCAAGAGAAATGCGCTGAAACCCCA 1740
1681 GCGGGCGCTGTAGTCCAGCTACTCAGAGGCTGAGCAAGAGAAATGCGCTGAAACCCCA 1740
1741 GGGGGCCGAGCTCAGTGAGCCGAGATTGGCCCACTGCACTCCAGCCTGGACGACAGTG 1800
1741 GGGGGCCGAGCTCAGTGAGCCGAGATTGGCCCACTGCACTCCAGCCTGGACGACAGTG 1800
1801 AGACTCCGTCTCAAAAAAAGAAAGAAAGAAAGCAAGCAAGCAAGTCTTTGGGGGGCAG 1860
1801 AGACTCCGTCTCAAAAAAAGAAAGAAAGAAAGCAAGCAAGCAAGTCTTTGGGGGGCAG 1860
1861 GGGGCGAGCAATAATTCTATACTTCCGGGATGCTGAGGGGGTTCATGGGAGGACCCCTG 1920
1861 GGGGCGAGCAATAATTCTATACTTCCGGGATGCTGAGGGGGTTCATGGGAGGACCCCTG 1920
1921 GCCTCTCTCTCCCAAGGATCCTCAACAGTGGTGTCAACAGGAAAAAATGCGAGCAATA 1980
1921 GCCTCTCTCTCCCAAGGATCCTCAACAGTGGTGTCAACAGGAAAAAATGCGAGCAATA 1980
1981 CGCTGAGGCTGTGGTCTTTCTGCTTTGAAAGGTCAGCTGTACTTAAAGGGACTGTTT 2040
1981 CGCTGAGGCTGTGGTCTTTCTGCTTTGAAAGGTCAGCTGTACTTAAAGGGACTGTTT 2040
2041 CAGCTCTCCCTGGGTGCTGCTTGGGACCCCTGCTGCAACCCCAACCACTCCCCCAACA 2100
2041 CAGCTCTCCCTGGGTGCTGCTTGGGACCCCTGCTGCAACCCCAACCACTCCCCCAACA 2100
2101 TCCTCTCTTTCCATCATATCCCCAGTATGAGCTTCCCAACTCCAGCCATAAGCTG 2160
2101 TCCTCTCTTTCCATCATATCCCCAGTATGAGCTTCCCAACTCCAGCCATAAGCTG 2160
2161 AATGTTTCTCTTAAAGATGGAGAAATCTTCTGCTCTCTGTCAGAAATTTGGGGAC 2220
2161 AATGTTTCTCTTAAAGATGGAGAAATCTTCTGCTCTCTGTCAGAAATTTGGGGAC 2220

2221 TGTGACTGGGATTTGTGGCTGGCTTAACTGCTGTGTGACCCAGACAGCCA 2280
2221 TGTGACTGGGATTTGTGGCTGGCTTAACTGCTGTGTGACCCAGACAGCCA 2280
2281 CTTCTCTCTCCCTAACTTTGGTTATGTCTTGGCAGCAAGTGAAGGCTGGAGCTAGGCGA 2340
2281 CTTCTCTCTCCCTAACTTTGGTTATGTCTTGGCAGCAAGTGAAGGCTGGAGCTAGGCGA 2340
2341 ACAGTTTGGATTTAGTGTGTTTGTAGTGTGGAATTTATTTTGTATATAAATCTTAT 2400
2341 ACAGTTTGGATTTAGTGTGTTTGTAGTGTGGAATTTATTTTGTATATAAATCTTAT 2400
2401 GTGTAACCCCAATATAGAACTAGATTAAAGAGGAGTCTCTCTGTTGAAAGGGGAGCTG 2460
2401 GTGTAACCCCAATATAGAACTAGATTAAAGAGGAGTCTCTCTGTTGAAAGGGGAGCTG 2460
2461 AGTACCTCTCTGGAACCTGAGGCACTCTGTAAGAAAGCAAACTGAAACCCAGTGGCTGGG 2520
2461 AGTACCTCTCTGGAACCTGAGGCACTCTGTAAGAAAGCAAACTGAAACCCAGTGGCTGGG 2520
2521 TCAGTGTCTCTTAAAGCAGTTTAAAGTGAGACCTGGAAACATTTGCTTTACCTT 2580
2521 TCAGTGTCTCTTAAAGCAGTTTAAAGTGAGACCTGGAAACATTTGCTTTACCTT 2580
2581 GAATAGATAGCTTTTATTTGTTGGTATATAAGAAATAAACTAACTATTAAACCTGAGAC 2640
2581 GAATAGATAGCTTTTATTTGTTGGTATATAAGAAATAAACTAACTATTAAACCTGAGAC 2640
2641 TTTACAGTGTGTTTATTTTATATAGTATATATAAAATTTTCTTTAGACATCAATTTTA 2700
2641 TTTACAGTGTGTTTATTTTATATAGTATATATAAAATTTTCTTTAGACATCAATTTTA 2700
2701 GGTAAAAAATAATTGATTAGAAATAATTTGCGAGGTCGACGCTCACCTGCAATCC 2760
2701 GGTAAAAAATAATTGATTAGAAATAATTTGCGAGGTCGACGCTCACCTGCAATCC 2760
2761 CAGGACTTTGGGAGCCGAGCGGCTGGATCACTCTGAGGTCAGGGGTTCAAGACAGAGCT 2820
2761 CAGGACTTTGGGAGCCGAGCGGCTGGATCACTCTGAGGTCAGGGGTTCAAGACAGAGCT 2820
2821 G 2821
2821 G 2821

RESULT 3
HSDJ139D8
LOCUS
DEFINITION
167078 bp DNA linear PRI 05-JUN-2003
Human DNA sequence from clone RPI-139D8 on chromosome 6p12.1-21.1
Contains genes for GUCA1A (guanylate cyclase activator 1A) and
GUCA1B (guanylate cyclase activator 1B) both from retina, the 3'
end of the tRep-132 gene for zinc finger transcription regulating
protein, the MRP510 gene for mitochondrial ribosomal protein S10,
the 5' end of the gene for a novel protein, the gene for a novel
protein and two CpG Islands, complete sequence.

ACCESSION
AL096814
VERSION
AL096814.26 GI:7159757
KEYWORDS
HTG; CpG Island; guanylate cyclase activator; GUCA1A; GUCA1B;
MRP510; tRep-132.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 167078)
Sehra,H.
Direct Submission
Submitted (05-JUN-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Mar 6, 2000 this sequence version replaced gi:7018706.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

<http://www.sanger.ac.uk/projects/C.elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr6>

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: humquery@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

RP1-139D8 is from the library RP1-1 constructed by the group of

Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2

This sequence is the entire insert of clone RP1-139D8 The true

right end of clone RP1-321B9 is at 5591 in this sequence.

FEATURES

source

1. 167078

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosomes="6"

/map="p12.1-21.1"

/clone="RP1-139D8"

/clone_lib="RP1-1"

35..49

repeat_region /note="3.8 copies 4 mer AGTT 23% conserved"

repeat_region 36..49

repeat_region /note="2.0 copies 7 mer GTTAGTT 28% conserved"

repeat_region complement(42..197)

repeat_region /note="L3 repeat: matches 1197..1345 of consensus"

repeat_region 62..72

repeat_region /note="2.2 copies 5 mer CTCTG 22% conserved"

repeat_region 132..151

repeat_region /note="2.5 copies 8 mer TTTGAAGA 40% conserved"

repeat_region 314..324

repeat_region /note="2.8 copies 4 mer TCAT 22% conserved"

repeat_region complement(327..425)

repeat_region /note="L2 repeat: matches 3210..3313 of consensus"

repeat_region complement(525..811)

repeat_region /note="L1MED repeat: matches 2400..2696 of consensus"

repeat_region 886..896

repeat_region /note="2.2 copies 5 mer TCATT 22% conserved"

repeat_region 966..1014

repeat_region /note="12.2 copies 4 mer GATA 91% conserved"

repeat_region 1054..1063

repeat_region /note="2.5 copies 4 mer AGCT 20% conserved"

repeat_region 1064..1077

repeat_region /note="7.0 copies 2 mer AT 28% conserved"

repeat_region 1079..1134

repeat_region /note="14.0 copies 4 mer TAGA 53% conserved"

repeat_region complement(1135..1245)

repeat_region /note="AluJb repeat: matches 195..288 of consensus"

repeat_region 1205..1215

repeat_region /note="3.7 copies 3 mer CCT 22% conserved"

repeat_region 1236..1274

repeat_region /note="7.8 copies 5 mer TTTTG 44% conserved"

repeat_region

1237..1273 /note="9.2 copies 4 mer TTTG 40% conserved"

repeat_region complement(1246..1553)

repeat_region /note="AluSx repeat: matches 5..312 of consensus"

repeat_region 1423..1682

repeat_region /note="1.9 copies 135 mer

CAACCACTTCAGCTCCCAAGTCTGGCATTACAGCATGACCACCGTCCCCAC

C 373% conserved"

repeat_region complement(1554..1692)

repeat_region /note="AluJb repeat: matches 1..195 of consensus"

repeat_region 1609

repeat_region complement(1695..1760)

repeat_region /note="L1MED repeat: matches 1010..1074 of consensus"

repeat_region 1761..2071

repeat_region /note="AluSx repeat: matches 1..312 of consensus"

repeat_region 1967..1979

repeat_region /note="2.2 copies 6 mer TGGAGG 26% conserved"

repeat_region 2042..2077

repeat_region /note="36.0 copies 1 mer A 36% conserved"

repeat_region 2045..2083

repeat_region /note="6.5 copies 6 mer AAAAAC 60% conserved"

repeat_region complement(2072..2752)

repeat_region /note="L1MED repeat: matches 318..1010 of consensus"

repeat_region 2144..2160

repeat_region /note="2.4 copies 7 mer TTTTGC 25% conserved"

repeat_region 2158..2168

repeat_region /note="2.8 copies 4 mer TTTA 22% conserved"

repeat_region 2709..2719

repeat_region /note="2.2 copies 5 mer GTGGA 22% conserved"

repeat_region 3048..3059

repeat_region /note="2.0 copies 6 mer ATTCT 24% conserved"

repeat_region 3137..3152

repeat_region /note="3.2 copies 5 mer CCTCT 32% conserved"

repeat_region 3378..3389

repeat_region /note="2.0 copies 6 mer CCTGAA 24% conserved"

repeat_region 3446..3464

repeat_region /note="2.4 copies 8 mer TCTCACC 29% conserved"

repeat_region 3463..3474

repeat_region /note="2.4 copies 5 mer CTTC 24% conserved"

repeat_region 3483..3494

repeat_region /note="2.0 copies 6 mer TTTTCC 24% conserved"

repeat_region 3569..3790

repeat_region /note="MIR repeat: matches 8..261 of consensus"

repeat_region 4077..4098

repeat_region /note="3.1 copies 7 mer CTGTGCC 26% conserved"

repeat_region 4103..4116

repeat_region /note="2.3 copies 6 mer TGGGSC 28% conserved"

repeat_region 4432..4473

repeat_region /note="L1MC1 repeat: matches 6287..6329 of consensus"

repeat_region complement(4587..4631)

repeat_region /note="L2 repeat: matches 3016..3054 of consensus"

repeat_region 5132..5142

repeat_region /note="2.2 copies 5 mer CCAGC 22% conserved"

repeat_region 5176..5194

repeat_region /note="3.8 copies 5 mer CCCAC 29% conserved"

repeat_region 5679..5688

repeat_region /note="2.0 copies 5 mer ATTAT 20% conserved"

repeat_region 5736..6034

repeat_region /note="AluSg repeat: matches 1..299 of consensus"

repeat_region 6016..6034

repeat_region /note="19.0 copies 1 mer A 38% conserved"

repeat_region 6052..6088

repeat_region /note="2.8 copies 13 mer CTACCCCTACTTC 56% conserved"

repeat_region 6092..6102

repeat_region /note="2.2 copies 5 mer TCTCT 22% conserved"

repeat_region complement(6113..6153)

repeat_region /note="MIR repeat: matches 33..73 of consensus"

repeat_region 6222..6237

repeat_region /note="2.0 copies 8 mer ATATCCTG 32% conserved"

repeat_region complement(6303..6388)

repeat_region /note="MIR3 repeat: matches 85..170 of consensus"

repeat_region 6380..6395

repeat_region /note="2.3 copies 7 mer CAGGCA 23% conserved"

repeat_region	6462. .6478	
repeat_region	/notes="4.2 copies 4 mer GGGT 27% conserved"	
repeat_region	6581. .6592	
repeat_region	/notes="2.0 copies 6 mer GGAGAG 24% conserved"	
repeat_region	7645. .7655	
repeat_region	/notes="5.5 copies 2 mer CT 22% conserved"	
repeat_region	7970. .7980	
repeat_region	/notes="2.2 copies 5 mer AACCA 22% conserved"	
repeat_region	8029. .8041	
repeat_region	/notes="2.6 copies 5 mer CAAAA 26% conserved"	
repeat_region	8047. .8057	
repeat_region	/note="2.2 copies 5 mer ACCCC 22% conserved"	
repeat_region	complement (8111. .8281)	
repeat_region	/notes="MIR repeat: matches 20. .192 of consensus"	
repeat_region	8146. .8156	
repeat_region	/notes="3.7 copies 3 mer TAT 22% conserved"	
repeat_region	8371. .8689	
repeat_region	/notes="AluY repeat: matches 1. .307 of consensus"	
repeat_region	8504. .8515	
repeat_region	/notes="12.0 copies 1 mer A 24% conserved"	
repeat_region	8665. .8709	
repeat_region	/note="9.0 copies 5 mer AAAAC 44% conserved"	
repeat_region	8666. .8712	
repeat_region	/notes="11.8 copies 4 mer AAC 42% conserved"	
repeat_region	8748. .8763	
repeat_region	/notes="16.0 copies 1 mer T 32% conserved"	
repeat_region	9005. .9122	
repeat_region	/note="MIR repeat: matches 50. .180 of consensus"	
Query Match	99.5%; Score 2805.8; DB 9; Length 167078;	
Best Local Similarity	99.9%; Pred. No. 0;	
Matches 2818; Conservative	0; Mismatches 2; Indels 1; Gaps 1;	
Qy	1	ATCCTAAGCGCACAGCTGGGAGGCGACACTGGGGAGTGGTCTGAGGATCTGGGT 60
Db	114296	ATCCTAAGCGCACAGCTGGGAGGCGACACTGGGGAGTGGTCTGAGGATCTGGGT 114355
Qy	61	CACCTGGGGTCAAGTGGAGCTTTTCAGGGTCTCCCGTCAATGAGCTGAGTTTCTTTGGC 120
Db	114356	CACCTGGGGTGAAGTGGAGCTTTTCAGGGTCTCCCGTCAATGAGCTGAGTTTCTTTGGC 114415
Qy	121	AGGGAATTTTACCAGCTGAAGAAAGCTGCGCGGAGAGCTACAACTGAGCAAGGCCAGC 180
Db	114416	AGGGAATTTTACCAGCTGAAGAAAGCTGCGCGGAGAGCTACAACTGAGCAAGGCCAGC 114475
Qy	181	TGCTCACACCGGAGGAGTGTGGAAGAGTCTTCTCTCTGTGTGAGTGAATGAGATG 240
Db	114476	TGCTCACACCGGAGGAGTGTGGAAGAGTCTTCTCTCTGTGTGAGTGAATGAGATG 114535
Qy	241	GTAAGAGGGCGAGAGTGGGAGAGTGTGTGCACATCTGCATCATCGCCACATTTCTGGCC 300
Db	114536	GTAAGAGGGCGAGAGTGGGAGAGTGTGTGCACATCTGCATCATCGCCACATTTCTGGCC 114595
Qy	301	GCAGCTCTTTGGGCAAGGCCCTCCACCTTTCCAAACCTCGGGGCTCTCATCTGTGAGAAGGC 360
Db	114596	GCAGCTCTTTGGGCAAGGCCCTCCACCTTTCCAAACCTCGGGGCTCTCATCTGTGAGAAGGC 114655
Qy	361	TGTGGAGAAGATGTATGAATCTAACAAAGGACTCATGAGCAGTGTGTTGTAGGAGTGAC 420
Db	114656	TGTGGAGAAGATGTATGAATCTAACAAAGGACTCATGAGCAGTGTGTTGTAGGAGTGAC 114715
Qy	421	TAAAGTCTTACAGGAGTTGCTCATGAGGCCAGGCACCGACGACGAGTGAAGATAGGAAC 480
Db	114716	TAAAGTCTTACAGGAGTTGCTCATGAGGCCAGGCACCGACGAGTGAAGATAGGAAC 114775
Qy	481	TTTGGAGTCAGGCAGGAGTGATATATTGAGCTTTCTCGCTAGTCTCAATTTCTCTCATC 540
Db	114776	TTTGGAGTCAGGCAGGAGTGATATATTGAGCTTTCTCGCTAGTCTCAATTTCTCTCATC 114835
Qy	541	TGGAAATGGGATATAATATGTTGGTTGAGAGGAATGAATAGGATTAATGTGTTTAAGAGC 600
Db	114836	TGGAAATGGGATATAATATGTTGGTTGAGAGGAATGAATAGGATTAATGTGTTTAAGAGC 114895

Db 1257 ATCCCAAGCGCACAGCTTGGGAAGCCAGCACTGGGAAGTGGTCTGGGGGATGTGGT 1316
QY 61 CACTGGGTGAAGGTGAGAGCTTTTCAAGGTCTCCCGTCAATGACGTGAGTTCCTTTGGC 120
Db 1317 CACTGGGTGAAGGTGAGAGCTTTTCAAGGTCTCCCGTCAATGACGTGAGTTCCTTTTTC 1376
QY 121 AGGGAATTTACAGCTGAAGAAAGCCTCCGCGAGAGCTACAAACTGAGCAAGCCACG 180
Db 1377 AGGGAATTTACAGCTGAAGAAAGCCTCCGCGAGAGCTACAGACTGAGCAAGCCACG 1436
QY 181 TGCTCACACCCGAGAGGTCTGGGACAGGATCTTCTCTGATGATGAGAATGAGATG 240
Db 1437 TGCTCACACCCGAGAGGTCTGGGACAGGATCTTCTCTGATGATGAGAATGAGATG 1496
QY 241 GTAAGAGGGGACAGATGGGAGAGTGTCTCCACTCTGCAATCATCGGCATTTCTGGCC 300
Db 1497 GTAAGAGGGGACAGATGGGAGAGTGTCTCCACTCTGCAATCATCGGCATTTCTGGCC 1556
QY 301 GCACGTCTTGGGCAAGGCCCTCCACCTTCCAAACCTGGGGTCTCATCTGTGAAAGGC 360
Db 1557 GCACGTCTTGGGCAAGGCCCTCCACCTTCCAAACCTGGGGTCTCATCTGTGAAAGGC 1616
QY 361 TGTGAGAGAGTGTCTATGAACCTAAACAAAGGGACTCATGAGCACGTGTTTGTAGGAGTGAC 420
Db 1617 TGTGAGAGAGTGTCTATGAACCTAAACAAAGGGACTCATGAGCACGTGTTTGTAGGAGTGAC 1676
QY 421 TAAAGTCTCAGAGATTGCTGATGGAGGCGAGGACGACAGAAATAGAAAGATAGGAAC 480
Db 1677 TAAAGTCTCAGAGATTGCTGATGGAGGCGAGGACGACAGAAATAGAAAGATAGGAAC 1736
QY 481 TTTTGAGTCAGCGAGGAGTGATATATTGAGCTTCTCGTCTAGTCTCAATTTCTCTCATC 540
Db 1737 TTTTGAGTCAGCGAGGAGTGATATATTGAGCTTCTCGTCTAGTCTCAATTTCTCTCATC 1796
QY 541 TGGAAATGGGGATTAATAGTGGTTGAGAGGAATGAATAGGATAATGTGTTTAAGAGC 600
Db 1797 TGGAAATGGGGATTAATAGTGGTTGAGAGGAATGAATAGGATAATGTGTTTAAGAGC 1856
QY 601 AGGCATAGGGTAGACCTTCATTCAGGCTGCTTGGGCTTTTCCCTCTGTAGGCCAAAGGCC 660
Db 1857 AGGCATAGGGTAGACCTTCATTCAGGCTGCTTGGGCTTTTCCCTCTGTAGGCCAAAGGCC 1916
QY 661 AGCCTCAGGCTATCTGGGAGAGAGCTGGCTTGAATACACACTTGAGCCCTCCAGCTC 720
Db 1917 AGCCTCAGGCTATCTGGGAGAGAGCTGGCTTGAATACACACTTGAGCCCTCCAGCTC 1976
QY 721 TCTCAGCTCCACCCAGCAATTTCCGTGTACCAATGCGCAAAAGTAAAACTTCAATTCATCA 780
Db 1977 TCTCAGCTCCACCCAGCAATTTCCGTGTACCAATGCGCAAAAGTAAAACTTCAATTCGTC 2036
QY 781 GCAAGAAAGCCCTTAAAGGTGGCAGGAGACTCTCGAGATTCAGACACCTGACAGCC 840
Db 2037 GCAAGAAAGCCCTTAAAGGTGGCAGGAGACTCTCGAGATTCAGACACCTGACAGCC 2096
QY 841 GCAAGCTTGAGCTCGAGCTCAGATAGTGGCATAGAGCTGTAGCGCATCTCGG 900
Db 2097 GCAAGCTTGAGCTCGAGCTCAGATAGTGGCATAGAGCTGTAGCGCATCTCGG 2156
QY 901 AGCGAGGTCTCTCTCCCTGCCCCAGACCAGGTCTCCCTTCTTCTACATGACCACCTC 960
Db 2157 AGCGAGGTCTCTCTCCCTGCCCCAGACCAGGTCTCCCTTCTTCTACATGACCACCTC 2216
QY 961 TCCTCCCTCTCTCAGGCGAGCTGTCTGAACAGATTTGTGAAGTGGCCGTCGGGA 1020
Db 2217 TCCTCCCTCTCTCAGGCGAGCTGTCTGAACAGATTTGTGAAGTGGCCGTCGGGA 2276
QY 1021 CAAAGTGGGTGAAGATGCTCAGATGACATGAATCCACAGCTGCTCGCTCAGCA 1080
Db 2277 CAAAGTGGGTGAAGATGCTCAGATGACATGAATCCACAGCTGCTCGCTCAGCA 2336
QY 1081 GAGACGGAAGTGCATGTTCTTGAGGAGTCTGGGGCCCCCTCCACGACTCCAGGCTCACC 1140

Db 2337 GAGACGGAAGTGCATGTTCTGAGGAGTCTGGGGCCCCCTCCACGACTCCAGGCTCACC 2396
QY 1141 CAGTTTTCAGGCTAGTAGGAGGCTCCCTGGCTCAGGCTGCTCATGCGCACTCTTCCCC 1200
Db 2397 CAGTTTTCAGGCTAGTAGGAGGCT--CCTGGCTCAGGCTGCTCATGCGCACTCTTCCCC 2454
QY 1201 TGGTGTTCACATCTCTGGCACCCCTGTGACGGCTGAGTGGGATGGGGAAGGGCTGCTG 1260
Db 2455 TGGTGTTCACATCTCTGGCACCCCTGTGACGGCTGAGTGGGATGGGGAAGGGCTGCTG 2514
QY 1261 GTTTGAAGTGGCCAAACAGGCTATGTCATTTTGGAGAGTCCCTGGGATGGTGAAGG 1320
Db 2515 GTTTGAAGTGGCCAAACAGGCTATGTCATTTTGGAGAGTCCCTGGGATGGTGAAGG 2574
QY 1321 AATTCAAGTACTTTTCTGTTTCAGCCGCTCTGGGAGACTGTGCTTTGGCTGGGTGTT 1380
Db 2575 AATTCAAGTACTTTTCTGTTTCAGCCGCTCTGGGAGACTGTGCTTTGGCTGGGTGTT 2634
QY 1381 GTGGGCTCCACAGTCTTCTGGGTGTTCTCAGTTTGGAAAGCAAGGCCAACTGAGGGGTGA 1440
Db 2635 GTGGGCTCCACAGTCTTCTGGGTGTTCTCAGTTTGGAAAGCAAGGCCAACTGAGGGGTGA 2694
QY 1441 GGGTCCCAACAGACCAATCAGAAATAGAAACACAAAGACTGGTAGGAGCGAGGGGTGGGA 1500
Db 2695 GGGTCCCAACAGACCAATCAGAAATAGAAACACAAAGACTGGTAGGAGCGAGGGGTGGGA 2754
QY 1501 GGGTCTTTCAGACTGAGAAAGGCGAGGAGTTGCCGGGCACGGTGGCTCACCGCTCTGAATC 1560
Db 2755 GGGTCTTTCAGACTGAGAAAGGCGAGGAGTTGCCGGGCACGGTGGCTCATGCTCTGAATC 2814
QY 1561 CCAGCACTTTTGGAGGCGGAGCGGCGAGATCACAGGTCAGGAGATCGAGACCAATCCTG 1620
Db 2815 CCAGCACTTTTGGAGGCGGAGCGGCGAGATCACAGGTCAGGAGATCGAGACCAATCCTG 2874
QY 1621 GCTAACACGGGTGAAACCCGCTCTCTACTAAAAATACAAAAAATCAGCCGGGTGAGGTG 1680
Db 2875 GCTAACACGGGTGAAAC--CTAATACTAATAAAAAATAAAAAATAAAGCCGTGTGTAGGTG 2932
QY 1681 GCGGGCGCTTGTAGTCCCGAGCTACTCAGGAGGCTGAGGCAAGAGATGGCGTGAACCCCA 1740
Db 2933 GCGGGCGCTTGTATCCAGCTACTCAGGAGGCTGAGGCAAGAGATTAGCGTGAACCCCA 2992
QY 1741 GCGGGC 1746
Db 2993 GCGGCC 2998

RESULT 5
HSM805690
LOCUS HSM805690 2288 bp mRNA linear PRI 16-JUN-2003
DEFINITION Homo sapiens mRNA; cDNA DKFZp686E1183 (from clone DKFZp686E1183); complete cds.
ACCESSION BX537393
VERSION BX537393.1 GI:31873287
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2288)
AUTHORS Koehrer,K., Beyer,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
TITLE Direct Submission
JOURNAL Submitted (16-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by BMFZ (Biomedical Research Center at the Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp686E1183) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further

information about the clone and the sequencing project is available
at <http://mips.gsf.de/proj/cdnaf/>.

FEATURES

source
1..2288
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /map="6p21.1"
 /clone="DKFZp686E1183"
 /tissue_type="human retina"
 /dh108_sites="SfiIA + SfiIB"
 /dev_stage="adult"
1..2288
 /gene="DKFZp686E1183"
137..739
 /gene="DKFZp686E1183"
 /notes="guanylate cyclase-activating protein 2"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAD97635.1"
 /db_xref="GI:31873288"
 /translation="MGQFSEWEAEAAAGRIDVAIEQWYKFPVMECPSTLPMHBFKR
PFKVTDDREASQYVSGMRAPDKNGDNTIDFLEYVALNLVLRGLEHKLKRTPLKID
KQNGCIDLKLELINVGGIYQKKACRRLEQTEQGOLLTPBEVDRIIFLLVDENGDDGQ
LSLNGFVSGARDKRWKMLQMDMNPSSLQAQRKSAHF"
2261..2266
 /polya_signal
2270
 /polya_site
 /gene="DKFZp686E1183"
 /gene="DKFZp686E1183"

ORIGIN

Query Match 58.0%; Score 1635.2; DB 9; Length 2288;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1659; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
QY 977 GGCAGCTGTCTCTGAACGAGTTTGTGAAGTGCCCGTGGGACAAAGTGGGTGATGAAG 1036
DB 611 GGCCAGCTGTCTCTGAACCGGGTTGTGAAGTGCCCGTGGGACAAAGTGGGTGATGAAG 670
QY 1037 ATGCTGCAGATGACATGAATCCAGCAGCTGGCTGCCTCAGCAGAGCGGAAAGTGCC 1096
DB 671 ATGCTGCAGATGACATGAATCCAGCAGCTGGCTGCCTCAGCAGAGCGGAAAGTGCC 730
QY 1097 ATGTTCTGAGGAGTCTGGGCGCCCTCCAGACTCCAGGCTCACCCAGGTTTCCAGGGTAG 1156
DB 731 ATGTTCTGAGGAGTCTGGGCGCCCTCCAGACTCCAGGCTCACCCAGGTTTCCAGGGTAG 790
QY 1157 TAGGAGGTCCCTGGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 1216
DB 791 TAGGAGGGT-CCCTGGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 849
QY 1217 GCACCCCTGTGCAGGGCTGAGTGGGATGGGAGGGGCTGTGGGTTGAAGTGGCCAA 1276
DB 850 GCACCCCTGTGCAGGGCTGAGTGGGATGGGAGGGGCTGTGGGTTGAAGTGGCCAA 909
QY 1277 CAGGGCATAGTCCATTTTGGAGGAGTCCCTGGGATGGTGAAGGGAATTCAGTTACTTTTC 1336
DB 910 CAGGGCATAGTCCATTTTGGAGGAGTCCCTGGGATGGTGAAGGGAATTCAGTTACTTTTC 969
QY 1337 CTGTTTCAGCCGCTCTCGGAGGACTGTGCTTTGGCTGGGTGGGTGGGTGGGTGGGTGGGTGG 1396
DB 970 CTGTTTCAGCCGCTCTCGGAGGACTGTGCTTTGGCTGGGTGGGTGGGTGGGTGGGTGGGTGG 1029
QY 1397 TTCTGGGTGTTCTCAGTTGGAGCAGAGCCCACTGAGGGGTGAGGGTCCCAAGCCCA 1456
DB 1030 TTCTGGGTGTTCTCAGTTGGAGCAGAGCCCACTGAGGGGTGAGGGTCCCAAGCCCA 1089
QY 1457 ATCAGAAATGAGAACACAAAGACTGGTAGGCGCAGGGGTGGAGGGTCTTCAGACTGAA 1516
DB 1090 ATCAGAAATGAGAACACAAAGACTGGTAGGCGCAGGGGTGGAGGGTCTTCAGACTGAA 1149
QY 1517 GAAAGGCGAGGAGTTGCCGGGACCGGTGGCTCAGCGCTGTAATCCAGCAGCTTTGGGAGG 1576

DB 1150 GAAAGGCGAGGAGTTGCCGGGACCGGTGGCTCAGCGCTGTAATCCAGCAGCTTTGGGAGG 1209
QY 1577 CCGAGCGGGCAGATCAGGAGTCAAGGATCGAGACCATCTCGGCTAAACCGGGGTGAA 1636
DB 1210 CCGAGCGGGCAGATCAGGAGTCAAGGATCGAGACCATCTCGGCTAAACCGGGGTGAA 1269
QY 1637 ACCCGCTCTACTATAAATAACAAAAATCAGCCGGGTGAGGTGGCGGGCGCTGTAGTC 1696
DB 1270 ACCCGCTCTACTATAAATAACAAAAATCAGCCGGGTGAGGTGGCGGGCGCTGTAGTC 1329
QY 1697 CCAGCTACTCAGAGGCTGAGGCAAGAGNATGGCTGAACCCAGCCGAGGGGCGGAGCCTACA 1756
DB 1330 CCAGCTACTCAGAGGCTGAGGCAAGAGNATGGCTGAACCCAGCCGAGGGGCGGAGCCTACA 1389
QY 1757 GTGAGCCGAGATTGGCCCACTGCACTCCAGCCTGACGACAGCTGAGACTCCGCTCTC-AAA 1815
DB 1390 GTGAGCCGAGATTGGCCCACTGCACTCCAGCCTGACGACAGCTGAGACTCCGCTCTCAAAA 1449
QY 1816 AAAAAAAGAAAGAAAGAAAGAGAGAGTTTGGGGGGCAGGGGCGAGCAATAAAT 1875
DB 1450 AAAAAAAGAAAGAAAGAAAGAGAGAGTTTGGGGGGCAGGGGCGAGCAATAAAT 1509
QY 1876 CTATAACTTCCGGGATGCTGAGGGGCTTCATGGGGAGGACCTTGGCTCTCTCTCCCA 1935
DB 1510 CTATAACTTCCGGGATGCTGAGGGGCTTCATGGGGAGGACCTTGGCTCTCTCTCCCA 1569
QY 1936 AGGCACTCTCACCACTGTGTCAACAGGAAAAATGCGCAAAATACGCTGCAAGGCTGTGG 1995
DB 1570 AGGCACTCTCACCACTGTGTCAACAGGAAAAATGCGCAAAATACGCTGCAAGGCTGTGG 1629
QY 1996 TCTTTCTGCTTTGAAAGGGTCACTGTACTTTAAAGGGACTGTCTTTCAGCTCTGCTGGGT 2055
DB 1630 TCTTTCTGCTTTGAAAGGGTCACTGTACTTTAAAGGGACTGTCTTTCAGCTCTGCTGGGT 1689
QY 2056 GCTGCTCTGGGACCCCTGCTGCCAACCCACACTCCCCCAACATCTCTCTCTTCCATC 2115
DB 1690 GCTGCTCTGGGACCCCTGCTGCCAACCCACACTCCCCCAACATCTCTCTCTTCCATC 1749
QY 2116 CATATCCCCAGTATGGACCTTCCCAACTCCAGCCATAAGCTGAATGTTTCTCTTTAA 2175
DB 1750 CATATCCCCAGTATGGACCTTCCCAACTCCAGCCATAAGCTGAATGTTTCTCTTTAA 1809
QY 2176 AGGATGGAGAAATCTCTGCTGCTGGCAAGAAATGGGGGACTGTGTGACTGGGATTG 2235
DB 1810 AGGATGGAGAAATCTCTGCTGCTGGCAAGAAATGGGGGACTGTGTGACTGGGATTG 1869
QY 2236 TGGGCTGGGCTTGGCTTCTAACTGCTGTGACCCCAAGACAGCCACTTCTCTCCCTAAC 2295
DB 1870 TGGGCTGGGCTTGGGCTTCTAACTGCTGTGACCCCAAGACAGCCACTTCTCTCCCTAAC 1929
QY 2296 CTTTGTATGCTTTCGCGACAGTGAAGTGGAGTCCGAGCTAGGCGAACAGTCTTGGATTAT 2355
DB 1930 CTTGTTATGCTTTCGCGACAGTGAAGTGGAGTCCGAGCTAGGCGAACAGTCTTGGATTAT 1989
QY 2356 TGTGTTTTTAGATGGAATTAATTTTTTGTATATAAACTTTATGTGTAAACCCCAATAT 2415
DB 1990 TGTGTTTTTAGATGGAATTAATTTTTTGTATATAAACTTTATGTGTAAACCCCAATAT 2049
QY 2416 AGAAACTAGATTAAAGGAGTCTCTGCTGTTGAAAGGGGAGCTGAGTACCCTCTGGAAC 2475
DB 2050 AGAAACTAGATTAAAGGAGTCTCTGCTGTTGAAAGGGGAGCTGAGTACCCTCTGGAAC 2109
QY 2476 TGGAGGCACTCTGAAAAAAGCAAACTGAAAAACAGTGCCTGGGTCACTGTTACTCTTA 2535
DB 2110 TGGAGGCACTCTGAAAAAAGCAAACTGAAAAACAGTGCCTGGGTCACTGTTACTCTTA 2169
QY 2536 TAAAGCAGTTTAAAGTGAAGCTGGAAAAAATTTGCTTTTACCTTTGAATAGATAGTTTT 2595
DB 2170 TAAAGCAGTTTAAAGTGAAGCTGGAAAAAATTTGCTTTTACCTTTGAATAGATAGTTTT 2229
QY 2596 TATGTTGGTATATAGAAATAAATAAATTAACCTTATTAACCCCTGAGA 2639

```
Db      2230 TATGTTGGTATATAGAAAATAAACTAACTATTATAAAGCTGAAA 2273

RESULT 6
HSM807325/c
LOCUS      HSM807325      2967 bp      mRNA      linear      PRI 30-AUG-2003
DEFINITION Homo sapiens mRNA; cDNA DKFZp686A2293 (from clone DKFZp686A2293).
ACCESSION BX647181
VERSION    BX647181.1 GI:34366209
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 2967)
AUTHORS    Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
CONSRSTM   The German Human cDNA Consortium
TITLE      Direct Submission
JOURNAL    Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
            Neuherberg, GERMANY
COMMENT    Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            This clone (DKFZp686A2293) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/proj/cDNA/.

FEATURES             source
     source
     1..2967
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /clone="DKFZp686A2293"
         /tissue type="human retina"
         /clone_lib="686 (synonym: hicc3). Vector pSport1_Sfi; host
         DH10B; sites SfiIA + SfiIB"
         /dev_stage="adult"
     polyA_site       2948
ORIGIN
Query Match      33.9%; Score 956.6; DB 9; Length 2967;
Best Local Similarity 99.3%; Pred. No. 1.9e-258;
Matches 972; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

QY 1532 GCCGGGCACGGTGGCTCAGCGCTGTAATCCCGACGACTTTGGAGGCCCGAGCGGCAGAT 1591
DB 996 GCCGGGCACGGTGGCTCAGCGCTGTAATCCCGACGACTTTGGAGGCCCGAGCGGCAGAT 937
QY 1592 CACGAGGTCAGGAGATCAGACCATCTCGGTAAACACGGGGTGAACCCCGTCTCTACTA 1651
DB 936 CACGAGGTCAGGAGATCAGACCATCTCGGTAAACACGGGGTGAACCCCGTCTCTACTA 877
QY 1652 AAAATACAAAAAATCAGCGGGTGAAGTGGCGGGCGCTGTAGTCCCGACTACTCAGGAG 1711
DB 876 AAAATACAAAAAATCAGCGGGTGAAGTGGCGGGCGCTGTAGTCCCGACTACTCAGGAG 817
QY 1712 GCTGAGGCAAGAGAAATGGCGTGAACCCCGAGGGGCCGAGCCCTACAGTGGCGGAGATTGC 1771
DB 816 GCTGAGGCAAGAGAAATGGCGTGAACCCCGAGGGGCCGAGCCCTACAGTGGCGGAGATTGC 757
QY 1772 GCCACTGCACTCCAGCTGGACGACGAGTGAAGTCCGTCCTC---AAAAAAGAGAA 1828
DB 756 GCCACTGCACTCCAGCTGGACGACGAGTGAAGTCCGTCCTC---AAAAAAGAGAA 697
QY 1829 AGAAAGAGAAAGCAGGAGTTTGGGGGCGAGGGGCGAGCAATAATCTATAACTTCCGG 1888
DB 696 AGAAAGAGAAAGCAGGAGTTTGGGGGCGAGGGGCGAGCAATAATCTATAACTTCCGG 637
QY 1889 GATGCTGAGGGCGCTTCATGGGAGGACCTGGCTCTCTCCCTCCCAAGGCATCCTACC 1948
|||||
```

```
Db      636 GATGCTGAGGGCGCTTCATGGGAGGAGCCCTCGGCTCCTCTCCCTCCCAAGGCATCCTACC 577
QY 1949 AGTGGTGTCAACAGAGAAATGGCAGCAAAATACGCTGCAGGCTGTGGTCTTTCTGCGCTTT 2008
DB 576 AGTGGTGTCAACAGAGAAATGGCAGCAAAATACGCTGCAGGCTGTGGTCTTTCTGCGCTTT 517
QY 2009 GAAAGGGTCACTGTACTTTAAAGGGAGCTGTTTTCAGCTCTGCTGGTGTGCTGCTCTGGGAC 2068
DB 516 GAAAGGGTCACTGTACTTTAAAGGGAGCTGTTTTCAGCTCTGCTGGTGTGCTGCTCTGGGAC 457
QY 2069 CCCCTGTGCGCAACCCAGCAGCTCCGCCAAACAATCTCTCTTTTCCATCCATATCCCCAGT 2128
DB 456 CCCCTGTGCGCAACCCAGCAGCTCCGCCAAACAATCTCTCTTTTCCATCCATATCCCCAGT 397
QY 2129 ATGACACCTTCCACAACATCCAGCAGCATAGCTGAATGTTTCTCTTTAAAGGATGGAGAAA 2188
DB 396 ATGACACCTTCCACAACATCCAGCAGCATAGCTGAATGTTTCTCTTTAAAGGATGGAGAAA 337
QY 2189 CTTCTGTCTGTCTCTGGCAAGAAATGGGGGAGCTGTTGACTGGGATTTGGGCTGGGCTTG 2248
DB 336 CTTCTGTCTGTCTCTGGCAAGAAATGGGGGAGCTGTTGACTGGGATTTGGGCTGGGCTTG 277
QY 2249 GCTTCTAACTGCTGTGACCCCAAGACGACCACTCTCTCCCTAACTTGGTTATGCT 2308
DB 276 GCTTCTAACTGCTGTGACCCCAAGACGACCACTCTCTCCCTAACTTGGTTATGCT 217
QY 2309 TGGCAGCACAGTGAAGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT 2368
DB 216 TGGCAGCACAGTGAAGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGT 157
QY 2369 GTGGAAATTTATTTTGTATATAAACTCTATGTGTAAACCCCAATATAGAAACTAGATTA 2428
DB 156 GTGGAAATTTATTTTGTATATAAACTCTATGTGTAAACCCCAATATAGAAACTAGATTA 97
QY 2429 AAAGGGAGTCTCTCTGGTTGAAAGGGAGCTGAGTACCTCTGGAACCTGGAGGACCTCT 2488
DB 96 AAAGGGAGTCTCTCTGGTTGAAAGGGAGCTGAGTACCTCTGGAACCTGGAGGACCTCT 37
QY 2489 GAAAAAGCAAACTGAAA 2507
DB 36 GAAAAAGCAAACTGAAA 18

RESULT 7
G30061/c
LOCUS      human STS SHGC-36091, sequence tagged site.
DEFINITION G30061
ACCESSION  G30061
VERSION     G30061.1 GI:1593612
KEYWORDS   STS; STS sequence; primer; sequence tagged site.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 407)
AUTHORS    Myers,R.M.
JOURNAL    Unpublished (1996)
COMMENT    Contact: Richard M. Myers
            Stanford Human Genome Center (SHGC)
            Stanford University School of Medicine
            Department of Genetics, M-344, Stanford, CA 94305, USA
            Tel: 4157259687
            Fax: 4157259689
            Email: myers@shgc.stanford.edu
            Primer A: CAAGGTAAAGCAAAATGTTTTC
            Primer B: AGGCACCTCTGAAAAAGCA
            STS size: 104
            PCR Profile:
                Initial incubation: 94 degrees C for 90 seconds
                Denaturation:      94 degrees C for 15 seconds
```

Annealing: 62 degrees C for 23 seconds
Polymerization: 72 degrees C for 30 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9600
Protocol:
Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
Tag Polymerase: 0.05 units/ul
Total Vol: 10 ul

Buffer: MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from H40682
-- Washington University/Merck EST sequence.

FEATURES
source

1..407
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="6"
38..141
38..60
complement(122..141)

ORIGIN

Query Match 13.3%; Score 374; DB 11; Length 407;
Best Local Similarity 98.0%; Pred. No. 6.3e-94;
Matches 399; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

QY 2214 GGGGACCTCTTCTACCTGGGATTTGGGCTTGGGCTTCTAAGTCTGTGTGACCCCAAG 2273

Db 407 GGGGGCTGTGACTGGGATTTGGGCTTGGGCTTCTAAGTCTGTGTGACCCCAAG 348

QY 2274 ACAGCCACTTCTCCTCCCTAACCTTGGTTATGTCTTGGCAGCAGTGTGAGCTGGAC 2333

Db 347 ACAGCCACTTCTCCTCCCTAACCTTGGTTATGTCTTGGCAGCAGTGTGAGCTGGAC 288

QY 2334 TAGGGCAACAGTTTGGATTTATTTTGTAGATGTGAATATTTTGTATATAA 2393

Db 287 TAGGGCAACAGTTTGGATTTATTTTGTAGATGTGAATATTTTGTATATAA 228

QY 2394 CTCTTATGTATACCCCAATAGAACTAGATTAAAGGGAGTCTCTGTGTGAAAGG 2453

Db 227 CTCTTATGTATACCCCAATAGAACTAGATTAAAGGGAGTCTCTGTGTGAAAGG 168

QY 2454 GGAG-CTGAGTACCTCTCGAACTGGAGGCACCTCTGAAAAAGCAAACTGAAAAACCACT 2512

Db 167 GGAGCTGAGTACCTCTCGAACTGGAGGCACCTCTGAAAAAGCAAACTGAAAAACCACT 108

QY 2513 GCCCTGGGTCACTGTACTCTTATAGA-CAGTTTAAAGTGAGACTTGAAAAACATTG 2571

Db 107 GCCCTGGGTCACTGTACTCTTATAGANCAGTTTAAAGTGAGCTGAAAAACATTG 48

QY 2572 CTTTACCTTGAATAGATAGGTTTATTTATTTTGTATATAAGAAATAA 2618

Db 47 CTTTACCTTGAATAGATAGGTTTATTTATTTTGTATATAAGCAATAA 1

RESULT 8

CQ491451/c
LOCUS CQ491451 1506 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 23318 from Patent WO0160860.
ACCESSION CQ491451
VERSION CQ491451.1 GI:41457070
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
TITLE Genes differentially expressed in human prostate cancer and their use
JOURNAL Patent: WO 0160860-A 23318 23-AUG-2001;
FEATURES Millennium Predictive Medicine, Inc. (US)
source Location/Qualifiers
1..1506
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 11.9%; Score 337; DB 6; Length 1506;
Best Local Similarity 99.7%; Pred. No. 2.1e-83;
Matches 348; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 977 GGCACGCTGTCTCTGAACGAGTTTGTGAAGTGGCCCTCGGACAAAGTGGGTGATGAAG 1036

Db 1325 GGCACGCTGTCTCTGAACGAGTTTGTGAAGTGGCCCTCGGACAAAGTGGGTGATGAAG 1266

QY 1037 ATGCTGCAGATGACATGAATCCAGCAGCTGGCTCGCTCAGCAGACGCGAAAGTGCC 1096

Db 1265 ATGCTGCAGATGACATGAATCCAGCAGCTGGCTCGCTCAGCAGACGCGAAAGTGCC 1206

QY 1097 ATGTTCTGAGGAGTCTGGGGCCCTCCAGACTCCAGGCTCACCAGGTTTCCAGGCTAG 1156

Db 1205 ATGTTCTGAGGAGTCTGGGGCCCTCCAGACTCCAGGCTCACCAGGTTTCCAGGCTAG 1146

QY 1157 TAGGAGGTCCCTCGCTCAGCTCTCATGCCCCACTCTTCCCTGGTGTGATTCCTG 1216

Db 1145 TAGGAGGT-CCCTGGCTCAGCTCTCATGCCCCACTCTTCCCTGGTGTGATTCCTG 1087

QY 1217 GCACCCCTCTGTCAGCGCTGAGTGGGATGGGGAAGGGCTGTGGTGTGAAAGTGCCAA 1276

Db 1086 GCACCCCTCTGTCAGCGCTGAGTGGGATGGGGAAGGGCTGTGGTGTGAAAGTGCCAA 1027

QY 1277 CAGGCGATAGTCCATTTTGGAGAGTCCCTGGGATGGTGAAGGAATTC 1325

Db 1026 CAGGCGATAGTCCATTTTGGAGAGTCCCTGGGATGGTGAAGGAATTC 978

RESULT 9

CQ497332/c
LOCUS CQ497332 1506 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 29199 from Patent WO0160860.
ACCESSION CQ497332
VERSION CQ497332.1 GI:41462968
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
TITLE Genes differentially expressed in human prostate cancer and their use
JOURNAL Patent: WO 0160860-A 29199 23-AUG-2001;
FEATURES Millennium Predictive Medicine, Inc. (US)
source Location/Qualifiers
1..1506
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 11.9%; Score 337; DB 6; Length 1506;
Best Local Similarity 99.7%; Pred. No. 2.1e-83;
Matches 348; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 977 GGCACGCTGTCTCTGAACGAGTTTGTGAAGTGGCCCTCGGACAAAGTGGGTGATGAAG 1036

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:

<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

```
----- Summary Statistics -----
Contig length: 191218
Phrap values in estimate: 190513
Average error rate (BCW-Phrap estimate): 3.84429e-05
Fraction of Phrap values less than 40: 0.0103405
Number of consensus changing edits: 15
Number of N's in consensus: 0
```

```
----- Consensus changing edits -----
Position Original+Context Edited+Context
5392 tctgactaa(n)attagttaga tctgactaa(a)attagttaga
72712 actgcaagct(n)tgctctccgg actgcaagct(c)tgctctccgg
72764 agtagctggg(g)ctacagccac agtagctggg(a)ctacagccac
72780 ggcaccacc(n)ccccaggcc ggcaccacc(a)ccccaggcc
72783 accacacc(c)ccaggctgg accacacc(c)ccaggctgg
72794 ccaggctgg(n)taaatcttcc ccaggctgg(c)taaatcttcc
72799 cctggntcaaa(a)ttttttttt cctggntcaaa(t)ttttttttt
72938 gccaccgtgc(n)cagcaaaaag gccaccgtgc(c)cagcaaaaag
108090 tttttttga(n)ggagctcac tttttttga(t)ggagctcac
108125 tggagtgccag(g)ggcacaatct tggagtgccag(t)ggcacaatct
108136 ggcacaatct(n)ggntcactgc ggcacaatct(c)ggntcactgc
159377 acaatcngg(n)tcactgcaac acaatcngg(c)tcactgcaac
181705 ttttgtgtgg(n)aaaaaaaaa ttttgtgtgg(a)aaaaaaaaa
181755 gctcttttta(n)cccttacatg gctcttttta(a)cccttacatg
```

----- Distribution of Quality < 40 Bases -----

```
1000 |
900 |
800 |
```

```
# bases
700 |
600 |
500 |
400 |
300 |
200 |
100 |
0 |
-----
5 10 15 20 25 30 35 40
Phrap Value Range
```

FEATURES

```
Version: 1.01 qxfo
Location/Qualifiers
1. .146530
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clones="RP11-438D14"
misc_feature
1. .6343
/note="overlaps bases 80628..86970 of clone AC016956"
repeat_region
694..795
/rpt_family="L2"
repeat_region
complement(796..1097)
/rpt_family="AluSp"
repeat_region
1098..1124
/rpt_family="L2"
STS
1417..1566
/standard_name="G07737"
/db_xref="dbSTS:13583"
repeat_region
complement(1615..1709)
/rpt_family="MIR"
repeat_region
complement(3095..3160)
/rpt_family="L2"
repeat_region
3317..3419
/rpt_family="MER104"
repeat_region
4052..4088
/rpt_family="AT-rich"
repeat_region
4220..4695
/rpt_family="MLT1J"
repeat_region
4922..4950
/rpt_family="AT-rich"
repeat_region
5402..5439
/rpt_family="(CAGA)n"
```

```
Query Match 9.9%; Score 278; DB 9; Length 146530;
Best Local Similarity 87.1%; Pred. No. 1.6e-66;
Matches 317; Conservative 0; Mismatches 45; Indels 2; Gaps 1;
QY 1488 GCGAGGGTGGGAGGGTGTTCAGACTGAAGAAAGGAGGAGTTGCCGGGACGGTGGCT 1547
DB 89495 GCGAGGTGCGAGAACAGGGTTATGTCAAGGTAATAATTTATCTAGCGCGGGCGGTGGCT 89436
QY 1548 CACGCTGTATCCAGCACCTTTGGGAGCGCGGCGGAGATCACCGGTCAGGAGAT 1607
DB 89435 CACGCTGTATCCAGCACCTTTGGGAGCGCGGCGGAGATCACCGGTCAGGAGAT 89376
QY 1608 CGAGACCATCTCGGTCTAACACGGGTGAACCCCGTCTCTACTAAAAATACAAAAATCA 1667
DB 89375 CGAGACCATCTCGGCTAACAC--GGTGAACCCCGTCTCTACTAAAAATACAAAAATTA 89318
QY 1668 GCCGGGTGAGTGGCGGGCGCTGTAGTCCAGCTACTCTAGAGGCTGAGGCAAGAGAT 1727
DB 89317 GCCGGGTGAGTGGCGGGCGCTGTAGTCCAGCTACTCTAGGAGGCTGAGGCAAGAGAT 89258
QY 1728 GCGGTGAACCCAGGGGCGCGGAGCTACAGTGAGCCGAGATTGGCCACTGCACCTCCAGC 1787
DB 89257 GCGGTGAACCCAGGGGCGCGGAGCTTCAGTGAGCCGAGATTGGCCCACTGCACCTCCAGC 89198
```

QY 1788 CTGGACGACAGTGGAGTCCGCTCTCAAAAAAAGAAAGAAAGAAAGGCGGAG 1847
 Db 89197 CTGGGCGACAGCGAGACTCCGCTCTCAAAAAAAGAAAGAAAGAAAGGCGGAG 89138
 QY 1848 TTTT 1851
 Db 89137 ATTT 89134

RESULT 12
 CR759924 74754 bp DNA linear HTG 24-SEP-2004
 LOCUS Homo sapiens chromosome 6 clone DAAP-388A15, *** SEQUENCING IN
 DEFINITION PROGRESS ***
 ACCESSION CR759924
 VERSION CR759924.2 GI:52673378
 KEYWORDS HTG; HTGS_PHASE2; HTGS_ACTIVEFIN.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 74754)
 AUTHORS Wood, J.
 TITLE Direct Submission
 JOURNAL Submitted (23-SEP-2004) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
 On Sep 24, 2004 this sequence version replaced gi:51965192.
 COMMENT ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk
 ----- Project Information
 Center project name: bAP388A15
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Chemistry: Dye-terminator; 100% of reads
 Consensus quality: 74142 bases at least Q40
 Consensus quality: 74380 bases at least Q30
 Consensus quality: 74580 bases at least Q20
 Insert size: 74754; sum-of-contigs
 Insert size: 83256; 2.6% error; agarose-fp
 Quality coverage: 6.02x in Q20 bases; sum-of-contigs Quality
 coverage: 5.69x in Q20 bases; agarose-fp

 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES
 source
 1..74754
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="6"
 /clone="DAAP-388A15"
 /clone_lib="APD"
 misc_feature 1..74754
 /note="assembly_fragment:00676"

ORIGIN
 Query Match 9.8%; Score 277.8; DB 2; Length 74754;
 Best Local Similarity 88.9%; Pred. No. 1.6e-66;
 Matches 312; Conservative 0; Mismatches 37; Indels 2; Gaps 1;
 QY 1515 AAGAAAGCAGGAGTTGCGGCGCAGGTGGCTGACGCTGTAATCCCGACACTTTGGGA 1574
 Db 24477 AAAAAAAGAAAGGCGGCGGTGCTACGCTGTAATCCCGACACTTTGGGA 24536
 QY 1575 GCGCGAGCGGCGGAGTTCAGAGGTTCAGAGATCGAGATCTCGGCTAAACACGGGGTG 1634
 Db 24537 GCGCGAGCGGCGGAGTTCAGAGGTTCAGAGATCGAGATCTCGGCTAAACAC--GGTG 24594

QY 1635 AAACCCCGTCTTCTACTAAAAATACAAAAATCAGCCGGGTGAGGTGGCGGGCGCTGTAG 1694
 Db 24595 AAACCCCGTCTTCTACTAAAAATACAAAAATTAGCCGGCGAGGTGGCGGGCGCTGTAG 24654
 QY 1695 TCCCGACTTACTCAGGAGGCTGAGGCAAGAGATGGCTGAACCCAGGGGGCGGACCTA 1754
 Db 24655 TCCTAGCTACTCGGAGGCTGAGGCGAGGAGATGGCTGAACCCAGGGGGCGGACCTG 24714
 QY 1755 CAGTGAGCCGAGATTGCGCCACTGCACCTCCAGCTTGGAGCAGCAGTGAGACTCCGCTCTCAA 1814
 Db 24715 CAGTGAGCCGAGATTGCGCCACTGCACCTCCAGCTTGGAGCAGCAGTGAGACTCCGCTCTCAA 24774
 QY 1815 AAAAAAAGAAAGAAAGAAAGAGGAGGAGTGGGGGGGCGACGGGGC 1865
 Db 24775 AAAAAAAGAAAGAAAGAAAGAAAGAAAGAAATTTGGGACAGATGTGGC 24825

RESULT 13
 CR788282 78827 bp DNA linear PRI 08-OCT-2004
 LOCUS Human DNA sequence from clone DAAP-200B17 on chromosome 6, complete
 DEFINITION sequence.
 ACCESSION CR788282
 VERSION CR788282.3 GI:54019828
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 78827)
 AUTHORS Palmer, S.
 TITLE Direct Submission
 JOURNAL Submitted (08-OCT-2004) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
 On Oct 9, 2004 this sequence version replaced gi:53748682.
 COMMENT ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr6/MHC
 DAAP-200B17 is from a APD human bac - B Lymphoblastoid Cell Line
 library VECTOR: pBelOBAC11.

FEATURES
 Location/Qualifiers
 1..78827
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6> DADB-23P23 is from a DNA-arts DBB human bac library VECTOR: pBelosAC11.

/chromosome="6"
/clone="DAAP-200B17"
/clone_lib="APD"

ORIGIN

Query Match 9.8%; Score 277.8; DB 9; Length 78827;
Best Local Similarity 88.9%; Pred. No. 1.7e-66;
Matches 312; Conservative 0; Mismatches 37; Indels 2; Gaps 1;

QY 1515 AAGAAAGCGAGGAGTTGCCCGGACAGGTGGCTACGGCTGTAAATCCAGCACTTTGGGA 1574
Db 51750 AAAAAAAAAAAAAAAAAAGCGCGCGGTGGCTACGGCTGTAAATCCAGCACTTTGGGA 51809

QY 1575 GCGGAGCGGCGAGATCACGAGGTCAGAGATCGAGACCATCTCTGGCTAAACACGGGCTG 1634
Db 51810 GCGGAGCGGCGGTGATCACGAGGTCAGAGATCGAGACCATCTCTGGCTAAACAC--GGTG 51867

QY 1635 AAACCCCGTCTCTACTAAAAATACAAAAATACAGCCGGGTGAGGTGGCGGCGCTGTAG 1694
Db 51868 AAACCCCGTCTCTACTAAAAATACAAAAATTAGCGGCGAGGTGGCGGCGCTGTAG 51927

QY 1695 TCCAGCTACTCAGAGGCTGAGGCAAGAGATGGCGTGAACCCAGGCGGCGAGCCTA 1754
Db 51928 TCCTAGCTACTCGGAGGCTGAGGCAAGAGATGGCGTGAACCCAGGCGGCGAGCCTG 51987

QY 1755 CAGTGAGCGGAGATGGCCCACTGCACCTCCAGCTTGGAGCAGAGTGAGACTCCGTCTCAA 1814
Db 51988 CAGTGAGCGGAGATGGCCCACTGCACCTCCAGCTTGGAGCAGAGACTCCGTCTCAA 52047

QY 1815 AAAAAAAAAAGAAAGAAAGGAGGAGGAGTTTGGGGGCGAGGGGCG 1865
Db 52048 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAATTTGGGACAGATGTGGC 52098

RESULT 14
CR388382/c
LOCUS Human DNA sequence from clone DADB-23P23 on chromosome 6, complete sequence.
DEFINITION CR388382.4 GI:47678787
ACCESSION HTG.
VERSION Homo sapiens (human)
KEYWORDS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
SOURCE Brown, J.
ORGANISM Direct Submission
Submitted (24-JUL-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On May 25, 2004 this sequence version replaced gi:47604289.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES
source
1.109067
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="DADB-23P23"
/clone_lib="DNA-arts-BAC.1-DBB.1"

ORIGIN

Query Match 9.8%; Score 277.8; DB 9; Length 109067;
Best Local Similarity 88.9%; Pred. No. 1.7e-66;
Matches 312; Conservative 0; Mismatches 37; Indels 2; Gaps 1;

QY 1515 AAGAAAGCGAGGAGTTGCCCGGACAGGTGGCTACGGCTGTAAATCCAGCACTTTGGGA 1574
Db 82323 AAAAAAAAAAAAAAAAAAGCGCGCGGTGGCTACGGCTGTAAATCCAGCACTTTGGGA 82264

QY 1575 GCGGAGCGGCGAGATCACGAGGTCAGAGATCGAGACCATCTCTGGCTAAACACGGGCTG 1634
Db 82263 GCGGAGCGGCGGTGATCACGAGGTCAGAGATCGAGACCATCTCTGGCTAAACAC--GGTG 82206

QY 1635 AAACCCCGTCTCTACTAAAAATACAAAAATACAGCCGGGTGAGGTGGCGGCGCTGTAG 1694
Db 82205 AAACCCCGTCTCTACTAAAAATACAAAAATTAGCGGCGAGGTGGCGGCGCTGTAG 82146

QY 1695 TCCAGCTACTCAGAGGCTGAGGCAAGAGATGGCGTGAACCCAGGCGGCGAGCCTA 1754
Db 82145 TCCTAGCTACTCGGAGGCTGAGGCAAGAGATGGCGTGAACCCAGGCGGAGCCTG 82086

QY 1755 CAGTGAGCGGAGATGGCGCACTGCACCTCCAGCTTGGAGCAGAGTGAGACTCCGTCTCAA 1814
Db 82085 CAGTGAGCGGAGATGGCGCACTGCACCTCCAGCTTGGAGCAGAGACTCCGTCTCAA 82026

QY 1815 AAAAAAAAAAGAAAGAAAGGAGGAGGAGTTTGGGGGCGAGGGGCG 1865
Db 82025 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAATTTGGGACAGATGTGGC 81975

RESULT 15
CR759926/c
LOCUS Homo sapiens chromosome 6 clone DAAP-269H15, *** SEQUENCING IN PROGRESS ***, 13 unordered pieces.
DEFINITION CR759926
ACCESSION CR759926.1 GI:51965193
VERSION HTG; HTGS_PHASE1.
KEYWORDS Homo sapiens (human)
SOURCE ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 148281)
Sims, S.
AUTHORS Direct Submission
TITLE Submitted (07-OCT-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
JOURNAL ----- Genome Center
COMMENT Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

CR759926 148281 bp DNA linear HTG 08-OCT-2004
Homo sapiens chromosome 6 clone DAAP-269H15, *** SEQUENCING IN PROGRESS ***, 13 unordered pieces.

CR759926
CR759926.1 GI:51965193
HTG; HTGS_PHASE1.
Homo sapiens (human)

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 148281)
Sims, S.

Direct Submission
Submitted (07-OCT-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 14:28:52 ; Search time 987.126 Seconds
(without alignments)
16917.381 Million cell updates/sec

Title: US-08-731-499-6
Perfect score: 2821
Sequence: 1 ATCCTAAGACGACACGCTG.....AGGGGTTCAAGACGACGCTG 2821

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:.*
1: Geneseqn1980s:.*
2: Geneseqn1990s:.*
3: Geneseqn2000s:.*
4: Geneseqn2001as:.*
5: Geneseqn2001bs:.*
6: Geneseqn2002as:.*
7: Geneseqn2002bs:.*
8: Geneseqn2003as:.*
9: Geneseqn2003bs:.*
10: Geneseqn2003cs:.*
11: Geneseqn2003ds:.*
12: Geneseqn2004as:.*
13: Geneseqn2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2808.6	99.6	2820	AAV09020	Homo sapi
2	337	11.9	1506	ABV29181	Human pro
3	337	11.9	1506	ABV23329	Human pro
4	325.8	11.5	425	AAI83630	Human pol
5	275	9.7	32152	AAK89020	Human dig
6	275	9.7	32152	AAK91534	Human dig
7	275	9.7	32152	AAI57791	Human col
8	275	9.7	32152	AAI57791	Human col
9	275	9.7	32152	ABG99968	Genomic D
10	275	9.7	32152	ADB32581	Human nov
11	275	9.7	32152	ADB93121	Human col
12	275	9.7	36933	AAK66362	Human imm
13	275	9.7	129042	ACN44674	Human gen
14	274.6	9.7	335	AAI99076	Human exc
15	274.6	9.7	335	AAI63426	Human kid
16	273.4	9.7	238484	ACN44210	Human gen
17	272.6	9.7	7995	AAI44507	Human LK
18	272.6	9.7	54378	ADQ59201	MSI-H car
19	271.8	9.6	45000	ADB99873	Partial 9
20	271.8	9.6	122186	AAK89560	Human his

C 21	271	9.6	3540	4	AAK72282	Human imm
C 22	270	9.6	810	4	AAI03110	Human rep
C 23	269.2	9.5	70469	11	ACN44014_3	Continuation (4 of
C 24	268.6	9.5	600	4	AAK71676	Human imm
C 25	266.8	9.5	15270	4	AAK74017	Human imm
C 26	266.8	9.5	15270	4	AAK85754	Human imm
C 27	266.4	9.4	51469	4	AAK78813	Human imm
C 28	266.4	9.4	51469	4	AAK70270	Human imm
C 29	266.4	9.4	51469	4	AAK69322	Human imm
C 30	266.4	9.4	334462	10	ADC24763	Human wil
C 31	263.6	9.3	816	4	AAI03109	Human rep
C 32	263.6	9.3	816	4	AAI03111	Human rep
C 33	263.6	9.3	101569	12	ADP13274	Renal cel
C 34	259.8	9.2	68495	13	ABD33553	Human can
C 35	259	9.2	110000	9	ADA13411_2	Continuation (3 of
C 36	259	9.2	110000	12	ADQ59446_2	Continuation (3 of
C 37	258.6	9.2	31304	4	AAI06868	Human rep
C 38	258.6	9.2	31304	4	ABA08089	Human ova
C 39	258.6	9.2	43936	11	ADL95498	Human MCM
C 40	258.6	9.2	44029	10	ADC85386	Human MCM
C 41	258.6	9.2	44030	9	ADA02906	Human MCM
C 42	258.6	9.2	44030	10	ADB72644	Human MCM
C 43	258.6	9.2	44030	12	ADM74501	Human car
C 44	257.4	9.1	15306	4	AAI07531	Human rep
C 45	257.2	9.1	57248	6	ABK83563	Human CDN

ALIGNMENTS

RESULT 1
AAV09020
ID AAV09020 standard; cDNA; 2820 BP.
AC AAV09020;
XX
XX
DT 21-JUL-1998 (first entry)
XX
XX Homo sapiens 20q13 amplicon GCAP transcript.
DE
XX
XX 20q13 amplicon; chromosome 20; tumour; detection: cyclic AMP; CAMP;
KW chromosomal abnormalities; probe; gene therapy; antisense inhibition;
KW treatment; age-related macular degeneration; retinitis pigmentata;
KW Leber's congenital amaurosis; biosynthesis;
KW guanine cyclase activating protein; ds.
XX
OS Homo sapiens.
XX
XX WO9802539-A1.
XX
XX PD 22-JAN-1998.
XX
XX PF 15-JUL-1997; 97WO-US012343.
XX
XX PR 15-JUL-1996; 96US-00680395.
PR 16-OCT-1996; 96US-00731499.
PR 17-JAN-1997; 97US-00785532.
XX
XX PA (REGC) UNIV CALIFORNIA.
XX
XX PI Gray JW, Collins CC, Hwang S, Godfrey T, Kowbel D, Rommens J;
XX WPI; 1998-110587/10.
XX
XX PT New sequences from the 20q13 amplicon - used for detecting chromosomal
PT abnormalities, particularly tumours, and for developing products for
PT treating diseases.
XX
XX PS Claim 1; Page 63; 91pp; English.
XX
XX CC The sequence is that of a cDNA sequence GCAP, which was isolated from the
CC 20q13 amplicon. It encodes a guanine cyclase activating protein which is
CC involved in the biosynthesis of cyclic AMP. Sequences from this gene can

be used in the treatment of retinal degeneration. It can be used as a probe for the detection of chromosomal abnormalities at 20q13. It and other sequences isolated from the 20q13 amplicon are consistently amplified in primary tumours. These sequences are useful as probes or as probe targets for monitoring the relative copy number of corresponding sequences from a biological sample such as tumour cells. The sequences can also be used in therapeutic applications for modulating the expression of the endogenous gene or the activity of the gene product. Examples of therapeutic approaches include antisense inhibition of gene expression, gene therapy, and monoclonal antibodies that specifically bind the gene products. The products can also be used in the treatment of other diseases, e.g. age-related macular degeneration, Leber's congenital amaurosis and retinitis pigmentata

Query Match	99.6%	Score 2808.6	DB %	Length 2820
Best Local Similarity	99.9%	Pred. No. 0		
Matches 2819	Conservative	1	Mismatches 0	Indels 1
Gaps	1			

Qy	1	ATCCTAAGACGCACAGCCTCTGGGAAGCCAGCACTCTGGGAAGTGGTCTGAGGATGTGGGT	60
Db	1	ATCCTTAAGACGCACAGCCTCTGGGAAGCCAGCACTGGGAAGTGGTCTGAGGATGTGGGT	60
Qy	61	CACCTGGGGTGAAGTGGAGCTTTCAAGGTCCTCCCGTCAATGCAGCTGAGTTTCTTTGGC	120
Db	61	CACCTGGGGTGAAGTGGAGCTTTCAAGGTCCTCCCGTCAATGCAGCTGAGTTTCTTTGGC	120
Qy	121	AGGGAATTTACAGCTGAAGAAAGCCTCTCGCGGAGAGCTACAAACTGAGCAAGGCCAGC	180
Db	121	AGGGAATTTACAGCTGAAGAAAGCCTCGCGGAGAGCTACAAACTGAGCAAGGCCAGC	180
Qy	181	TGCTCACACCGAGAGGTCTGTGGACAGGATTTCTCTCTGTGTGATGAGAAATGAGATG	240
Db	181	TGCTCACACCGAGAGGTCTGTGGACAGGATTTCTCTCTGTGTGATGAGAAATGAGATG	240
Qy	241	GTAAGAGGGCAGAGATGGGAGAGTGTCTCCACTCTGCATCATCGCCACCTTTCTGGCC	300
Db	241	GTAAGAGGGCAGAGATGGGAGAGTGTCTCCACTCTGCATCATCGCCACCTTTCTGGCC	300
Qy	301	GCAGTCTCTTGGGCAAGGCCCTTCCACTTCCAACCTCTGGGTCCTCATCTGTGAAAGGC	360
Db	301	GCAGTCTCTTGGGCAAGGCCCTTCCACTTCCAACCTCTGGGTCCTCATCTGTGAAAGGC	360
Qy	361	TGTGGAGAGATGTCAATGAATTAACAAAGGACTCATGAGCAGCTGTTGTAGGAGTGC	420
Db	361	TGTGGAGAGATGTCAATGAATTAACAAAGGACTCATGAGCAGCTGTTGTAGGAGTGC	420
Qy	421	TAAAGTCTCAGGAGTTGCTGATGGAGGCCAGGCACGCAATAGAAAGAAATAGGAAC	480
Db	421	TAAAGTCTCAGGAGTTGCTGATGGAGGCCAGGCACGCAATAGAAAGAAATAGGAAC	480
Qy	481	TTTGGAGTCAGGCAGGGAGTGATATATTGAGCTTCTCGTCTAGTCTCAATTTCTCTCATC	540
Db	481	TTTGGAGTCAGGCAGGGAGTGATATATTGAGCTTCTCGTCTAGTCTCAATTTCTCTCATC	540
Qy	541	TGGAAATGGGGATTAATATAGTGTGGAGGAATGAATAGGATAATGTGTTTAAGAGC	600
Db	541	TGGAAATGGGGATTAATATAGTGTGGAGGAATGAATAGGATAATGTGTTTAAGAGC	600
Qy	601	AGGCATGGGTAGACCTCCATTTCAGGCTGTCTGGGCTTTTCCCTCTGTAGCCCAAGGCC	660
Db	601	AGGCATGGGTAGACCTCCATTTCAGGCTGTCTGGGCTTTTCCCTCTGTAGCCCAAGGCC	660
Qy	661	AGCCTCAGGGCTATGTGGGGAGAGCTGGCTTGGAAATACACTTGAGCCCTCCAGCTC	720
Db	661	AGCCTCAGGGCTATGTGGGGAGAGCTGGCTTGGAAATACACTTGAGCCCTCCAGCTC	720
Qy	721	TCTCAGCTCCACCCAGCATTTCCGTGGTACCAATGCGCAAAAGTAACCTCAATTCAATCA	780
Db	721	TCTCAGCTCCACCCAGCATTTCCGTGGTACCAATGCGCAAAAGTAACCTCAATTCAATCA	780
Qy	781	GCAAGAAAGCCCTTTAAAGGTGGCAGGAGACTCTGGGAGATTTCAGACACCTTCAGACGC	840

Db 1861 GGGGAGCAATATTTCTATATACTTCGGGATGCTGAGGGGGTTCATGGGAGGACCCTG 1920
Qy 1921 GCCTCCTCTCCCAAGGCATCTCACAGTGGTGTCAACAGGAAATGGCAGCAATA 1980
Db 1921 GCCTCCTCTCCCAAGGCATCTCACAGTGGTGTCAACAGGAAATGGCAGCAATA 1980
Qy 1981 CGCTGAGGCTGTGCTTTCTGCTTTGAAAGGTCAGCTGACTTAAAGGAGCTGTTT 2040
Db 1981 CGCTGAGGCTGTGCTTTCTGCTTTGAAAGGTCAGCTGACTTAAAGGAGCTGTTT 2040
Qy 2041 CAGCTCTGCTGGGTGCTCTCTGGGACCCCTGCTGCCAACCCACCTCTCCCAACAA 2100
Db 2041 CAGCTCTGCTGGGTGCTCTCTGGGACCCCTGCTGCCAACCCACCTCTCCCAACAA 2100
Qy 2101 TCCTCTCTTTCCATCCATCCCCAGTATGCACTTCCACACTCCAGCCATAGCTG 2160
Db 2101 TCCTCTCTTTCCATCCATCCCCAGTATGCACTTCCACACTCCAGCCATAGCTG 2160
Qy 2161 AATGTTCTCTTAAAGGATGAGAAACTTCTGCTGCTCTGCAAGAAATGGGGAC 2220
Db 2161 AATGTTCTCTTAAAGGATGAGAAACTTCTGCTGCTCTGCAAGAAATGGGGAC 2220
Qy 2221 TGTGACTGGGATGTGGGCTGGCTTCTAACTGCTGTGTGACCAAGACAGCCA 2280
Db 2221 TGTGACTGGGATGTGGGCTGGCTTCTAACTGCTGTGTGACCAAGACAGCCA 2280
Qy 2281 CTTCTCTCCCTAACCTTGTATGCTTGGCAGCAGTGTGAGCGAGTGGGACTAGCGA 2340
Db 2281 CTTCTCTCCCTAACCTTGTATGCTTGGCAGCAGTGTGAGCGAGTGGGACTAGCGA 2340
Qy 2341 ACAGTTTGGATATGCTTTTATGATGTGAAATATTTTGTATATAAATCTTAT 2400
Db 2341 ACAGTTTGGATATGCTTTTATGATGTGAAATATTTTGTATATAAATCTTAT 2400
Qy 2401 GTGTAACCCCAATATAGAACTAGATTAAGAGGAGTCTCTGCTGTTGAAAGGGAGCTG 2460
Db 2401 GTGTAACCCCAATATAGAACTAGATTAAGAGGAGTCTCTGCTGTTGAAAGGGAGCTG 2460
Qy 2461 AGTACCCCTCTGAACTGGAGGACCTCTGAAAGGACCAACTGAAAGGAGGAGGCTGGG 2520
Db 2461 AGTACCCCTCTGAACTGGAGGACCTCTGAAAGGACCAACTGAAAGGAGGAGGCTGGG 2520
Qy 2521 TCAGTGTACTCTATAGACAGTTTAAAGTGAACCTGGAAGGACCAATTTGCTTTACCTT 2580
Db 2521 TCAGTGTACTCTATAGACAGTTTAAAGTGAACCTGGAAGGACCAATTTGCTTTACCTT 2580
Qy 2581 GAATAGATAGTGTGTTTATGCTGTATAGAAATAAAGTAACTAATTAACCTGAGAC 2640
Db 2581 GAATAGATAGTGTGTTTATGCTGTATAGAAATAAAGTAACTAATTAACCTGAGAC 2640
Qy 2641 TTTACAGGTGTGTTATTTTATATAGTATAGTATATAAATTTCTTTAGACATCAATTTTA 2700
Db 2641 TTTACAGGTGTGTTATTTTATATAGTATAGTATATAAATTTCTTTAGACATCAATTTTA 2700
Qy 2701 GGTAAAAAATATGATTAGAAAAATATTTGGCAGGTGAGCAGCTCAACCTGCAATCC 2760
Db 2701 GGTAAAAAATATGATTAGAAAAATATTTGGCAGGTGAGCAGCTCAACCTGCAATCC 2760
Qy 2761 CAGGACTTTGGAGGCGGAGGGGTGATCACCTGAGGTCAAGGTTTCAAGCAGGCT 2820
Db 2761 CAGGACTTTGGAGGCGGAGGGGTGATCACCTGAGGTCAAGGTTTCAAGCAGGCT 2820
Qy 2821 G 2821
Db 2820 G 2820

RESULT 2

ABV29181/c

ID ABV29181 standard; cDNA; 1506 BP.

XX

AC ABV29181;

XX

DT 16-SEP-2002 (first entry)
XX Human prostate expression marker cDNA 29172.
DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX Homo sapiens.
OS
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX Schlegel R, Endege WO, Monahan JE;
PI
XX WPI; 2001-662795/76.
XX
DR Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
PT
XX
PS Claim 1; Page 6211-6212; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 1506 BP; 348 A; 444 C; 421 G; 290 T; 0 U; 3 Other;
Query Match 11.9%; Score 337; DB 5; Length 1506;
Best Local Similarity 99.7%; Pred. No. 1.4e-76;
Matches 348; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 977 GGCACGCTGTCTCTCAACGAGTTTGTGAAGTGGCCCTCGGACAGTGGGTGATGAAG 1036
Db 1325 GGCACGCTGTCTCTCAACGAGTTTGTGAAGTGGCCCTCGGACAGTGGGTGATGAAG 1266
Qy 1037 ATGCTGCAGATGGACATGAATCCAGCAGCTGGCTCGCTCAGCAGACGCGAAAGTGCC 1096
Db 1265 ATGCTGCAGATGGACATGAATCCAGCAGCTGGCTCGCTCAGCAGACGCGAAAGTGCC 1206
Qy 1097 ATGTTCTGAGGAGTCTGGGGCCCTCCACGACTCCAGGCTCACCAGGTTTCCAGGGTAG 1156
Db 1205 ATGTTCTGAGGAGTCTGGGGCCCTCCACGACTCCAGGCTCACCAGGTTTCCAGGGTAG 1146
Qy 1157 TAGGAGGTCCCTGGCTCAGCTGCTCATGCGCACTCTTCCCTGGTGTGACTTCCTG 1216
Db 1145 TAGGAGGT-CCCTGGCTCAGCTGCTCATGCGCACTCTTCCCTGGTGTGACTTCCTG 1087
Qy 1217 GCACCCCTGTGTCAGGCTGAGTGGGGATGGGGAAGGGGTGCTGGTGTGAAAGTGCCAA 1276
Db 1086 GCACCCCTGTGTCAGGCTGAGTGGGGATGGGGAAGGGGTGCTGGTGTGAAAGTGCCAA 1027

QY	1277	CAGGCATAGTCCATTTGGAGAGTCCCTGGGATGCTGAAGGGAATTC	1325	Db	1205	ATGTTCTGAGAGTCTGGGGCCCTCCACGACTCAGCCAGGTTTCCAGGCTAG	1146
Db	1026	CAGGCATAGTCCATTTGGAGAGTCCCTGGGATGCTGAAGGGAATTC	978	QY	1157	TAGAGGGTCCCTGGCTCAGCCTGCTCATGCCACACTTCTCCCTGGTGTGACTTCCTG	1216
RESULT 3				Db	1145	TAGGAGGGT-CCCTGGCTCAGCCTGCTCATGCCACACTTCTCCCTGGTGTGACTTCCTG	1087
ID	ABV23329/c			QY	1217	GCACCCCTGTGCAGGGCTGAGTGGGATGGGAGGGCTGCTGGGTTTGAAGTGCCAA	1276
XX	ABV23329			Db	1086	GCACCCCTGTGCAGGGCTGAGTGGGATGGGAGGGCTGCTGGGTTTGAAGTGCCAA	1027
XX	16-SEP-2002	(first entry)		QY	1277	CAGGCATAGTCCATTTGGAGAGTCCCTGGGATGCTGAAGGGAATTC	1325
XX	Human prostate expression marker	cdna 23320.		Db	1026	CAGGCATAGTCCATTTGGAGAGTCCCTGGGATGCTGAAGGGAATTC	978
KW	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;			RESULT 4			
KW	pharmacogenomic marker; gene; ss.			AAI83630			
OS	Homo sapiens.			ID	AAI83630	standard; cDNA; 425 BP.	
PN	WO200160860-A2.			XX	XX		
XX	23-AUG-2001.			AC	AAI83630;		
XX	20-FEB-2001; 2001WO-US005171.			XX	06-NOV-2001	(first entry)	
XX	17-FEB-2000; 2000US-0183319P.			XX	Human polynucleotide	SEQ ID NO 3690.	
PR	16-MAR-2000; 2000US-0189862P.			XX	Human; cytokine; cell proliferation; cell differentiation; gene therapy;		
PR	25-MAY-2000; 2000US-0207454P.			KW	vaccine, peptide therapy; stem cell growth factor; haematopoiesis;		
PR	09-JUN-2000; 2000US-0211314P.			KW	tissue growth factor; immunomodulatory; cancer; leukaemia;		
PR	18-JUL-2000; 2000US-0219007P.			KW	nervous system disorders; arthritis; inflammation; ss.		
PR	13-DEC-2000; 2000US-0255281P.			OS	Homo sapiens.		
XX	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.			XX	WO200164835-A2.		
PA	Schlegel R, Endege WO, Monahan JE;			PN	07-SEP-2001.		
PI	WPI; 2001-662795/76.			XX	26-FEB-2001; 2001WO-US004927.		
XX	Novel isolated nucleic acid molecule associated with cancerous state of			XX	28-FEB-2000; 2000US-00515126.		
PT	prostate cells and correlating with presence of prostate cancer, useful			PR	18-MAY-2000; 2000US-00577409.		
PT	for detecting presence of prostate cancer, stage of prostate cancer.			XX	(HYSE-) HYSEQ INC.		
PS	Claim 1; Page 4227; 11750pp; English.			PA	Tang YT, Liu C, Drmanac RT;		
XX	The invention relates to an isolated nucleic acid molecule (I) comprising			PI	WPI; 2001-514838/56.		
CC	a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the			XX	P-PSDB; AAC03699.		
CC	specification or its complement. (I) is useful for: (a) assessing whether			DR	Isolated nucleic acids and polypeptides, useful for preventing diagnosing		
CC	a patient is afflicted with prostate cancer; (b) monitoring the			DR	and treating e.g. leukemia, inflammation and immune disorders.		
CC	progression of prostate cancer in a patient; (c) assessing the efficacy			XX	Claim 1; SEQ ID NO 3690; 1399pp + Sequence Listing; English.		
CC	of a test compound to inhibit prostate cancer in a patient; (d) assessing			PT	The invention relates to human polynucleotides (AAI79941-AAI93841) and		
CC	the efficacy of a therapy for inhibiting prostate cancer in a patient;			XX	the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to		
CC	(e) selecting a composition for inhibiting prostate cancer in a patient;			CC	cytokine, cell proliferation or cell differentiation or which may induce		
CC	(f) assessing the prostate cell carcinogenic potential of a compound; (g)			CC	production of other cytokines in other cell populations. The		
CC	determining whether prostate cancer has metastasized in a patient; (h)			CC	polynucleotides and polypeptides are useful in gene therapy, vaccines or		
CC	assessing the aggressiveness or indolence of prostate cancer in a patient			CC	peptide therapy. The polypeptides have various cytokine-like activities,		
CC	; (I) is also useful as a pharmacodynamic or pharmacogenomic marker			CC	e.g. stem cell growth factor activity, haematopoiesis regulating		
XX	Sequence 1506 BP; 348 A; 444 C; 421 G; 290 T; 0 U; 3 Other;			CC	activity, tissue growth factor activity, immunomodulatory activity and		
QY	Query Match	11.9%; Score 337; DB 5; Length 1506;		CC	activin/inhibin activity and may be useful in the diagnosis and/or		
XX	Best Local Similarity	99.7%; Pred. No. 1.4e-76;		CC	treatment of cancer, leukaemia, nervous system disorders, arthritis and		
XX	Matches 348; Conservative	0; Mismatches 0; Indels 1; Gaps 1;		CC	inflammation. Note: The sequence data for this patent did not form part		
QY	977	GGCCAGCTGTCTCTGAACGAGTTTCTGAAGTGCCTCGGACCAAGTGGTGATGAAG	1036	CC	of the printed specification, but was obtained in electronic format		
Db	1325	GGCCAGCTGTCTCTGAACGAGTTTCTGAAGTGCCTCGGACCAAGTGGTGATGAAG	1266	CC	directly from WIPO at ftp.wipo.int/pub/published_pat_sequences		
QY	1037	ATGCTGCAGATGGACATGAATCCAGCAGCTGGCTCGCTCAGACAGCGAAGTCCC	1096	XX	Sequence 425 BP; 125 A; 111 C; 102 G; 86 T; 0 U; 1 Other;		
Db	1265	ATGCTGCAGATGGACATGAATCCAGCAGCTGGCTCGCTCAGACAGCGAAGTCCC	1206	Query Match	11.5%; Score 325.8; DB 4; Length 425;		
QY	1097	ATGTTCTGAGGAGTCTGGGGCCCTCCACGACTCCAGGCTCACCCAGGTTTCCAGGCTAG	1156	Best Local Similarity	86.1%; Pred. No. 6.3e-74;		
				Matches 360; Conservative	0; Mismatches 58; Indels 0; Gaps 0;		

PR	18-AUG-2000;	2000US-02262799;
PR	22-AUG-2000;	2000US-0226681P.
PR	22-AUG-2000;	2000US-02266868P.
PR	23-AUG-2000;	2000US-0227182P.
PR	23-AUG-2000;	2000US-02277009P.
PR	30-AUG-2000;	2000US-0228924P.
PR	01-SEP-2000;	2000US-0229287P.
PR	01-SEP-2000;	2000US-0229343P.
PR	01-SEP-2000;	2000US-0229344P.
PR	01-SEP-2000;	2000US-0229345P.
PR	05-SEP-2000;	2000US-0229509P.
PR	05-SEP-2000;	2000US-0229513P.
PR	06-SEP-2000;	2000US-0230437P.
PR	06-SEP-2000;	2000US-0230438P.
PR	08-SEP-2000;	2000US-0231242P.
PR	08-SEP-2000;	2000US-0231243P.
PR	08-SEP-2000;	2000US-0231244P.
PR	08-SEP-2000;	2000US-0231413P.
PR	08-SEP-2000;	2000US-0231414P.
PR	08-SEP-2000;	2000US-0232080P.
PR	08-SEP-2000;	2000US-0232081P.
PR	13-SEP-2000;	2000US-0231968P.
PR	14-SEP-2000;	2000US-0232397P.
PR	14-SEP-2000;	2000US-0232398P.
PR	14-SEP-2000;	2000US-0232399P.
PR	14-SEP-2000;	2000US-0232400P.
PR	14-SEP-2000;	2000US-0232401P.
PR	14-SEP-2000;	2000US-0233063P.
PR	14-SEP-2000;	2000US-0233064P.
PR	14-SEP-2000;	2000US-0233065P.
PR	21-SEP-2000;	2000US-0234223P.
PR	21-SEP-2000;	2000US-0234224P.
PR	23-SEP-2000;	2000US-0236327P.
PR	23-SEP-2000;	2000US-0236367P.
PR	25-SEP-2000;	2000US-0234997P.
PR	25-SEP-2000;	2000US-0234998P.
PR	26-SEP-2000;	2000US-0235484P.
PR	27-SEP-2000;	2000US-0235634P.
PR	27-SEP-2000;	2000US-0235636P.
PR	29-SEP-2000;	2000US-0236327P.
PR	29-SEP-2000;	2000US-0236367P.
PR	02-OCT-2000;	2000US-0236802P.
PR	02-OCT-2000;	2000US-0237037P.
PR	02-OCT-2000;	2000US-0237038P.
PR	02-OCT-2000;	2000US-0237039P.
PR	02-OCT-2000;	2000US-0237040P.
PR	13-OCT-2000;	2000US-0239935P.
PR	13-OCT-2000;	2000US-0239937P.
PR	20-OCT-2000;	2000US-0241808P.
PR	20-OCT-2000;	2000US-0241809P.
PR	20-OCT-2000;	2000US-0241821P.
PR	20-OCT-2000;	2000US-0241826P.
PR	01-NOV-2000;	2000US-024617P.
PR	08-NOV-2000;	2000US-0246474P.
PR	08-NOV-2000;	2000US-0246475P.
PR	08-NOV-2000;	2000US-0246476P.
PR	08-NOV-2000;	2000US-0246477P.
PR	08-NOV-2000;	2000US-0246478P.
PR	08-NOV-2000;	2000US-0241785P.
PR	20-OCT-2000;	2000US-0241786P.
PR	20-OCT-2000;	2000US-0241787P.
PR	20-OCT-2000;	2000US-0241808P.
PR	20-OCT-2000;	2000US-0241809P.
PR	20-OCT-2000;	2000US-0241826P.
PR	01-NOV-2000;	2000US-024617P.
PR	08-NOV-2000;	2000US-0246474P.
PR	08-NOV-2000;	2000US-0246475P.
PR	08-NOV-2000;	2000US-0246476P.
PR	08-NOV-2000;	2000US-0246477P.
PR	08-NOV-2000;	2000US-0246478P.
PR	08-NOV-2000;	2000US-024523P.
PR	08-NOV-2000;	2000US-024532P.
PR	08-NOV-2000;	2000US-0246609P.
PR	08-NOV-2000;	2000US-0246610P.
PR	08-NOV-2000;	2000US-0246611P.

PR	08-NOV-2000;	2000US-0245613P	
PR	17-NOV-2000;	2000US-0249207P	
PR	17-NOV-2000;	2000US-0249208P	
PR	17-NOV-2000;	2000US-0249209P	
PR	17-NOV-2000;	2000US-0249210P	
PR	17-NOV-2000;	2000US-0249211P	
PR	17-NOV-2000;	2000US-0249212P	
PR	17-NOV-2000;	2000US-0249213P	
PR	17-NOV-2000;	2000US-0249214P	
PR	17-NOV-2000;	2000US-0249215P	
PR	17-NOV-2000;	2000US-0249216P	
PR	17-NOV-2000;	2000US-0249217P	
PR	17-NOV-2000;	2000US-0249218P	
PR	17-NOV-2000;	2000US-0249244P	
PR	17-NOV-2000;	2000US-0249245P	
PR	17-NOV-2000;	2000US-0249264P	
PR	17-NOV-2000;	2000US-0249265P	
PR	17-NOV-2000;	2000US-0249297P	
PR	17-NOV-2000;	2000US-0249299P	
PR	17-NOV-2000;	2000US-0249300P	
PR	01-DEC-2000;	2000US-0250160P	
PR	01-DEC-2000;	2000US-0250391P	
PR	05-DEC-2000;	2000US-0251030P	
PR	05-DEC-2000;	2000US-0251988P	
PR	05-DEC-2000;	2000US-0256719P	
PR	06-DEC-2000;	2000US-0251479P	
PR	08-DEC-2000;	2000US-0251856P	
PR	08-DEC-2000;	2000US-0251868P	
PR	08-DEC-2000;	2000US-0251869P	
PR	08-DEC-2000;	2000US-0251988P	
PR	08-DEC-2000;	2000US-0251990P	
PR	11-DEC-2000;	2000US-0254907P	
PR	05-JAN-2001;	2001US-0259678P	
XX	(HUMA-) HUMAN GENOME SCI INC.		
XX			
XX	Rosen CA, Barash SC, Ruben		
PI			
XX	WPI; 2001-502630/55.		
DR			
XX			
XX			
PT	Polynucleotides encoding digest		
PT	diagnosing, treating, prevent		
PT	digestive system, particul		
XX			
XX	Disclosure; SEQ ID NO 5110; 9		
XX			
CC	The present invention provide		
CC	number of human digestive sys		
CC	diagnosis, treatment and prev		
CC	including cancer, Meckel's di		
CC	infections, appendicitis, Hir		
CC	ulcerative colitis. The prese		
CC	encoding a digestive system a		
XX			
SQ	Sequence 32152 BP; 8332 A; 73		
	Query Match	9.7%;	
	Best Local Similarity	88.1%;	
	Matches	311; Conservative	0
Qy	1486 GAGGCAGGGGTGGGAGGGTGG		
Db	18284 GAGGCGGGTGGATCATGAAG		
Qy	1546 CTCACGCGTGTAAATCCAGC		
Db	18224 CTCACGCGTGTAAATCCAGC		
Qy	1606 ATCGAGACCATCTGGCTAA		
Db	18164 ATCGAGACCATCTGGCTAA		
Qy	1666 CAGCCGGGTGAGGTGGCGG		

```
Db 18106 TAGCGGCGAGGTGGCGGCGCTGTAGTCCACGCTACTCGGAGGCTGAGGCAGGAGA 18047
Qy 1726 ATGGCGTGAACCCAGCGGGCGAGCCCTACAGTGAAGCGAGATTGCGGCACCTGCACCTCCA 1785
Db 18046 ATGGCGTGAACCCAGCGGGCGAGCCCTGCACTGAGCGGAGATTGCGGCACCTGCACCTCCA 17987
Qy 1786 GCCTGGAGCAGCAGTCAGACTCGTCTCAAAAAAAAAAAAAAGAAAGAAA 1838
Db 17986 GCCTGGCGGACAGCGAGACTCGTCTCAAAAAAAAAAAAAAGAAAAGAAA 17934

RESULT 7
AAI57791/C
ID AAI57791 standard; DNA; 32152 BP.
XX AC AAI57791;
XX DT 19-OCT-2001 (first entry)
XX DE Human colorectal cancer antigen coding sequence SEQ ID NO: 328.
XX KW Human; colorectal cancer; colorectal cancer antigen; gene therapy; ds.
XX OS Homo sapiens.
XX PN WO200155350-A1.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001350.
XX PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
```


PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241828P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.

PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465567/50.
XX
DR Isolated polypeptide for treating, preventing and/ or prognosing
XX disorders related to the colon including colon cancers and also for
XX testing and detection e.g. diagnosis.
XX
PS Disclosure; SEQ ID NO 518; 562pp; English.
XX
CC The present invention relates to the isolation of novel human colon
CC associated polypeptides (AAU22468-AAU22701), and the cDNA and genomic
CC sequences encoding for them. The sequences of the invention are useful in
CC the diagnosis, treatment, prevention and/or prognosis of disorders of the
CC colon including colon cancer, congenital abnormalities (e.g. atresia and
CC stenosis), bacterial and viral infections, inflammatory bowel disease
CC (IBD), neoplastic cell disorders (e.g. polyps and adenomas, intestinal
CC inflammatory disorders, colitis, colonic inflammation, diarrhoea and
CC dysentery, malabsorption syndromes (e.g. lactose intolerance), intestinal
CC obstruction and sigmoid diseases. The polynucleotide sequences of the
CC invention can also be used in gene therapy. AAS39582-AAS40060 represent
CC DNA sequences encoding for the novel human colon associated polypeptides
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 32152 BP; 9229 A; 7256 C; 7335 G; 8332 T; 0 U; 0 Other;
Query Match 9.7%; Score 275; DB 5; Length 32152;
Best Local Similarity 88.1%; Pred. No. 7.5e-60;
Matches 311; Conservative 0; Mismatches 40; Indels 2; Gaps 1;
QY 1486 GAGCGGGGTGGGGGTGTGAGACTGAAGAAAGCAGGAGTTGCCGGCAGCGTGG 1545
DB 13869 GAGCGGGTGGATCATGAAGTGAAGAGATCAAGACCATCTCGTTGGCGCGCGTGG 13928
QY 1546 CTCACGCTGTAAATCCAGCACTTTGGGAGGCCGAGGCGGAGATCACGAGGTGAGAG 1605
DB 13929 CTCACGCTGTAAATCCAGCACTTTGGGAGGCCGAGGCGGAGATCACGAGGTGAGAG 13988
QY 1606 ATCGAGACCATCTCGGTAAACACGGGGTGAACCCCGTCTTACTTAAATAATACAAAAAT 1665
DB 13989 ATCGAGACCATCTCGGTAAACAC--GGTGAACCCCGTCTTACTTAAATAATACAAAAAT 14046
QY 1666 CAGCGGGTGGAGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAGGTGAGGCAAGAGA 1725
DB 14047 TAGCCGGCGAGGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAGGTGAGGCAAGAGA 14106
QY 1726 ATGGCGTGAACCCAGAGGGGCGAGCCCTACAGTGAAGCCAGATTGCGCACTGCCA 1785
DB 14107 ATGGCGTGAACCCAGAGGGGCGAGCCCTGAGTGAAGCCAGATTGCGCACTGCCA 14166
QY 1786 GCCTGGACGACGTGAGACTCCGTCTCAAAAAAAGAAAAAGAAAAAGAAA 1838

Db 14167 GCCTGGGACAGCGAGACTCGTCTCAAAAAAAAAAAAAAAAAAAGA 14219

RESULT 9
ABS99968/c
ID ABS99968 standard; DNA; 32152 BP.
XX
AC ABS99968;
XX
DT 18-DEC-2002 (first entry)
XX
DE Genomic DNA #172 encoding human colorectal cancer related protein.
XX
KW Human; colorectal cancer related protein; colon; rectum;
KW colorectal cancer metastasis; gastrointestinal disorder; cytostatic;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN US2002119919-A1.
XX
PD 29-AUG-2002.
XX
PF 17-JAN-2001; 2001US-00764855.
XX
PR 31-JAN-2000; 2000US-0179065P.
XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX WPI; 2002-731367/79.
XX
PT New colorectal cancer polypeptide for diagnosing, prognosing, preventing,
PT and treating immune, hyperproliferative, liver, kidney, reproductive
PT disorders and for identifying modulators of therapeutic use.
XX
PS Disclosure; SEQ ID NO 328; 183pp; English.
XX
CC The present invention relates to the isolation of novel human colorectal
CC cancer related proteins, and polynucleotide sequences encoding them. The
CC sequences of the invention are useful in the diagnosis, treatment,
CC prevention and/or prognosis of the colon and/or rectum, including
CC colorectal cancer, colorectal cancer metastases, and gastrointestinal
CC disorders such as dysphagia, peptic esophagitis, gastric reflux, and
CC irritable bowel syndrome, and peritoneal diseases. The invention also
CC describes antibodies that bind colorectal cancer related proteins,
CC vectors, host cells, and recombinant and synthetic methods for producing
CC human colorectal cancer related polynucleotides, polypeptides, and/or
CC antibodies. ABS99979-ABS99974 represent genomic sequences encoding human
CC colorectal cancer related proteins. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the USPTO web site at
XX seqdata.uspto.gov/psipdsIDentry.html
XX
SQ Sequence 32152 BP; 8332 A; 7335 C; 7256 G; 9229 T; 0 U; 0 Other;

Query Match 9.7%; Score 275; DB 6; Length 32152;
Best Local Similarity 88.1%; Pred. No. 7.5e-60;
Matches 311; Conservative 0; Mismatches 40; Indels 2; Gaps 1;

QY 1486 GAGGCGGGTGGGAGGGTGTGAGACTGAACAAAGGCGAGGTTGCCGGGCACGGTGG 1545
Db 18284 GAGGCGGGTGGGATCATGATGAAGATCAAGACCATCTCTGGTGGCGGGCGGGTGG 18225

QY 1546 CTCACGCTGTATCCACGACTTTGGGAGGCGGCGGAGATCAGAGTTCAGGAG 1605
Db 18224 CTCACGCTGTATCCACGACTTTGGGAGGCGGCGGAGATCAGAGTTCAGGAG 18165

QY 1606 ATCGAGACCATCTCGGTAAACCGGGGTGAACCCCGTCTTACTAAAAATACAAAAAT 1665
|||||

Db 18164 ATCGAGACCATCTCGGTAAACAC--GGTGAAACCCCGTCTTACTAAAAATACAAAAAT 18107

QY 1666 CAGCCGGGTGAGGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAGGCTGAGGCAAGAGA 1725
|||||

Db 18106 TAGCCGGGCGAGGTGGCGGGCGCTGTAGTCCAGCTACTCGGAGGCTGAGGCAAGAGA 18047
|||||

QY 1726 ATGGCGTGAAACCCAGGGGGCGGAGCCTACAGTACGCGAGATTGGCCACTGCCTCCA 1785
|||||

Db 18046 ATGGCGTGAAACCCAGGGGGCGGAGCCTGCGAGTGGCCGAGATTGGCCACTGCCTCCA 17987
|||||

QY 1786 GCCTGGGACAGCTGAGACTCGTCTCAAAAAAAAAAAAAAAAAAAGAAA 1838
|||||

Db 17986 GCCTGGGACAGCTGAGACTCGTCTCAAAAAAAAAAAAAAAAAAAGAAA 17934
|||||

RESULT 10
ADB32581
ID ADB32581 standard; DNA; 32152 BP.
XX
AC ADB32581;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human novel colon related polypeptide DNA SEQ ID NO 518.
XX
KW gene therapy; cancer; liver disorder; hepatitis; neural disorder;
KW Alzheimer's disease; human; colon; ds.
XX
OS Homo sapiens.
XX
PN US2003050231-A1.
XX
PD 13-MAR-2003.
XX
PF 17-JAN-2001; 2001US-00764872.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 28-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227003P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.

```
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249266P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0259678P.
XX
FA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-625420/59.
XX
PT New nucleic acid molecule, useful for preparing a medicament for
PT preventing, treating or ameliorating a medical condition e.g. cancer,
PT liver disorders or neural disorders.
XX
PS Disclosure; SEQ ID NO 518; 216pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule encoding a
CC polypeptide. The nucleic acid is useful for preparing a medicament for
CC preventing, treating or ameliorating a medical condition e.g. cancer,
CC liver disorders such as hepatitis or neural disorders such as Alzheimer's
CC disease. The present sequence represents a human novel colon related
CC polypeptide DNA. Note: The sequence data for this patent did not form
CC part of the printed specification but was obtained in electronic format
CC directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030050231.
XX
SQ Sequence 32152 BP; 9229 A; 7256 C; 7335 G; 8332 T; 0 U; 0 Other;

Query Match          9.7%; Score 275; DB 9; Length 32152;
Best Local Similarity 88.1%; Pred. No. 7.5e-60;
Matches 311; Conservative 0; Mismatches 40; Indels 2; Gaps 1;

QY 1486 GAGGACGGGTGGGAGGGTGTGAGACTGAGAAAGACAGAGTTCGCGGCGACGGTGG 1545
DB 13869 GAGGCGGGTGGATCATGAAGTGAAGAGATCAAGACCATCTGTTGGCCGGCGGGTGG 13928
QY 1546 CTCACGCCTGTAATCCAGCAGCACTTTGGGAGGCCGAGGCGGCAGATCAAGAGTTCAGGAG 1605
DB 13929 CTCACGCCTGTAATCCAGCAGCACTTTGGGAGGCCGAGGCGGCAGATCAAGAGTTCAGGAG 13988
QY 1606 ATCGAGACCATCTGGCTTAACACGGGGTGAACCCCGTCTCTACTAAAAATACAAAAAT 1665
DB 13989 ATCGAGACCATCTGGCTTAACAC--GGTGAACCCCGTCTCTACTAAAAATACAAAAAT 14046
QY 1666 CAGCGCGGTGAGGTGGCGGCGCTGTAGTCCAGTACTCAGGAGGTGAGGCAAGAGA 1725
DB 14047 TAGCCGGCGAGGTGGCGGCGCTGTAGTCCAGTACTCAGGAGGTGAGGCAAGAGA 14106
QY 1726 ATGGCGTGAACCCAGGCGGCGGAGCCTTACAGTGAGCCGAGATTGCGCCACTGCACTCCA 1785
```

Db 14107 ATGGCGTGAACCCAGGGGCGGAGCCTGTCAGTGGCCGAGATTGGCCACTGCACTCCA 14166
QV 1786 GCGTCGACGACAGTGGAGCTCCGCTCTCAAAAAAAGAAAAAGAAAGAAAAGAAA 1838
Db 14167 GCGTCGCGACAGCGAGCTCCGCTCTCAAAAAAAGAAAAAAGAAAAAAGAAAAGA 14219
RESULT 11
ADB93121/c
ID ADB93121 standard; DNA; 32152 BP.
XX ADB93121;
AC ADB93121;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human colorectal cancer related polypeptide DNA #172.
KW ds; gene; human; colorectal cancer; antigen; gene therapy;
KW gastrointestinal disorder; inflammatory disease; infection; cancer;
KW intestinal neoplasm; small intestine carcinoma tumour;
KW small intestine non-Hodgkin's lymphoma; small bowel lymphoma; ulcer;
KW peptic ulcer; Bruton's disease; X linked infantile agammaglobulinaemia;
KW severe combined immunodeficiency; DiGeorge anomaly;
KW hyperproliferative disorder; acute lymphoblastic leukaemia;
KW acute lymphocytic leukaemia; urinary system disorder; cortical necrosis;
KW kidney infarction; cardiovascular disorder; carcinoma heart disease;
KW arrhythmia; respiratory disorder; non-allergic rhinitis; sinusitis;
KW musculoskeletal system disorder; Albers-Schonberg disease;
KW Marfan's syndrome; neurological disease; phenylketonuria;
KW Wernicke's encephalopathy; Alzheimer's disease; endocrine disorder;
KW Grave's disease; Cushing's syndrome; reproductive system disorder;
KW prostatitis; benign prostatic hypertrophy; benign prostatic hyperplasia;
KW thrombosis; atherosclerosis; myocardial infarction; ischaemic attack.
OS Homo sapiens.
XX
XX US2003054420-A1.
XX
XX 20-MAR-2003.
XX
XX 11-FEB-2002; 2002US-00072349.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 07-JUN-2000; 2000US-0205515P.
PR 28-JUN-2000; 2000US-0209467P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0227009P.
PR 01-SEP-2000; 2000US-022824P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 05-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235583P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-023637P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2000US-0259678P.
PR 17-JAN-2001; 2001US-00764855.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-708345/67.
XX
XX Novel colorectal cancer antigen useful for treating, preventing,
XX diagnosing and/or prognosing gastrointestinal disorders, infections,
XX cancers such as intestinal neoplasms, ulcers.
XX
XX Disclosure; SEQ ID NO 328; 179pp; English.
XX
XX The invention relates to a colorectal cancer antigen. The antigen is
XX useful for chromosome identification, chromosome mapping, radiation
XX hybrid mapping or gene therapy, or as hybridisation probes for
XX differential identification of the tissues or cell types present in a
XX biological sample. The antigen is useful for treating, preventing,
XX diagnosing and/or prognosing gastrointestinal disorders, including
XX inflammatory diseases and/or conditions, infections, cancers (e.g.
XX intestinal neoplasms (carcinoid tumour of the small intestine, non-
XX Hodgkin's lymphoma of the small intestine, small bowel lymphoma) and
XX ulcers (e.g. peptic ulcers). The antigen and its nucleic acid are useful
XX to provide immunological probes for differential identification of the
XX tissue. The antigen and its nucleic acid are useful for treating,
XX preventing, diagnosing and/or prognosing diseases, disorders and/or
XX conditions of the immune system e.g. Bruton's disease, X linked infantile
XX agammaglobulinaemia, severe combined immunodeficiencies, DiGeorge
XX anomaly, etc. The antigen and its nucleic acid is useful for treating,
XX preventing and/or diagnosing hyperproliferative disorders (e.g. acute
XX lymphoblastic leukaemia, acute lymphocytic leukaemia, etc), urinary

Query Match 9.7%; Score 275; DB 10; Length 32152;
Best Local Similarity 88.1%; Pred. No. 7, 5e-60;
Matches 311; Conservative 0; Mismatches 40; Indels 2; Gaps 1;
Qy 1486 GAGGCGGGTGGAGGTGTGAGACTGAGAAAGCAGGAGTTGCCGGGCACCGTGG 1545
Db 18284 GAGGCGGGTGGATCATGAAGTGAAGAGATCAAGACCATCCTGGTTGGCCGGCGGTGG 18225

Qy 1546 CTCACGCTGTAAATCCAGCACTTTTGGAGGCCGAGGGCGGCGAGATCACGAGGTCAGGAG 1605
Db 18224 CTCACGCTGTAAATCCAGCACTTTTGGAGGCCGAGGGCGGCGGATCACGAGGTCAGGAG 18165
Qy 1606 ATCGAGACCATCTCTGGGTTAAACACGGGGTGAAACCCCGTCTCTACTTAAAAATACAAAAAT 1665
Db 18164 ATCGAGACCATCTCTGGGTTAAACAC--GGTGAACCCCGTCTCTACTTAAAAATACAAAAAT 18107
Qy 1666 CAGCCGGGTGAGGTGGCGGGCGCTGTAGTCCGAGCTACTCAGGAGGCTGAGGCAAGAGA 1725
Db 18106 TAGCCGGCGGAGGTGGCGGGCGCTGTAGTCCGAGCTACTCAGGAGGCTGAGGCAAGAGA 18047
Qy 1726 ATGGCGTGAACCCAGCGGGCGGAGCCCTACAGTGAGCGGAGATTGCGGCCTGCACTCCA 1785
Db 18046 ATGGCGTGAACCCAGCGGGCGGAGCCCTGCAGTGAGCGGAGATTGCGGCCTGCACTCCA 17987
Qy 1786 GCCTGGACGACAGTGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAGAAAGAA 1838
Db 17986 GCCTGGCGACAGCGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAGAA 17934
RESULT 12
AAK66362
ID AAK66362 standard; DNA; 36933 BP.
XX
XX AAK66362;
XX
XX 06-NOV-2001 (first entry)
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:21174.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225575P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0223287P.
 PR 01-SEP-2000; 2000US-0223433P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 06-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0233080P.
 PR 08-SEP-2000; 2000US-0233081P.
 PR 14-SEP-2000; 2000US-0233401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234977P.
 PR 26-SEP-2000; 2000US-0234990P.
 PR 27-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239335P.
 PR 13-OCT-2000; 2000US-0239337P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246509P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-483426/52.

XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 useful for preventing, diagnosing and/or treating cancers and metastasis.

PS Disclosure; SEQ ID NO 21174; 3071pp + Sequence Listing; English.

XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients' own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention
 XX Sequence 36933 BP; 10500 A; 8526 C; 8341 G; 9566 T; 0 U; 0 Other;

Query Match 9.7%; Score 275; DB 4; Length 36933;

Best Local Similarity 88.1%; Pred. No. 9e-60;

Matches 311; Conservative 0; Mismatches 40; Indels 2; Gaps 1;

OY 1486 GAGCGGGTGGGAGGGTGTGACACTGACAAAGAGCAGGAGTTGCCGGCCACGGTGG 1545

Db 13869 GAGCGGGTGGGATCATGATGAAGTGAAGATCAAGACCATCTCTGTTGGCGGGCGGGTGG 13928

OY 1546 CTCACGCCTGTAATCCACGACCTTTGGGAGGCCGCGGAGGCAGATCAGAGGTGAGGAG 1605

Db 13929 CTCAGCCCTGTAATCCAGCACTTTGGAGGCGCGAGGCTCAGAGGTCAGGAG 13988
Qy 1606 ATCGAGACATCTGGCTAAACACGGGGTGAACCCCGCTCTACTATAAAATACAAAAAT 1665
Db 13989 ATCGAGACATCTGGCTAAACAC--GGTGAACCCCGCTCTACTATAAAATACAAAAAT 14046
Qy 1666 CAGCCGGGTGAGGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAGGCTGAGGCAGAGA 1725
Db 14047 TAGCCGGCGAGGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAGGCTGAGGCAGAGA 14106
Qy 1726 ATGGCGTGAACCCACAGGGGGCGAGCCCTACAGTGAGCCGAGATTGCGCACTGCACTCCA 1785
Db 14107 ATGGCGTGAACCCACAGGGGGCGAGCCCTGAGTGAGCCGAGATTGCGCACTGCACTCCA 14166
Qy 1786 GCCTGGAGCAGATGAGACTCGCTCTCAAAAAAAGAAAAAGAAAAA 1838
Db 14167 GCCTGGCGCAGCAGCAGACTCGCTCTCAAAAAAAGAAAAAAGAAAAA 14219

RESULT 13
ACN44674
ID ACN44674 standard; DNA; 129042 BP.

AC ACN44674;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human genomic sequence hCG27276.
XX
KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX
OS Homo sapiens.
XX
FN WO2003073826-A2.
XX
PD 12-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-328604/31.
XX
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.

XX Claim 1; SEQ ID NO 1240; Opp; English.
XX
CC The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published
XX
SQ Sequence 129042 BP; 35555 A; 29165 C; 28632 G; 35690 T; 0 U; 0 Other;

Query Match 9.7%; Score 275; DB 11; Length 129042;
Best Local Similarity 88.1%; Pred. No. 1.5e-59;
Matches 311; Conservative 0; Mismatches 40; Indels 2; Gaps 1;

Qy 1486 GAGCAGGGTGGAGGGTGTGTGAGACTGAAGAAAAAGCAGAGTTGCCGCGCACCGGTGG 1545
Db 123352 GAGCGGGTGGATCATGAGTGAAGAGATCAAGACCATCTCTGGTTGGCGGGCGGTGG 123411
Qy 1546 CTCACGCTGTAATCCAGCACTTTGGAGGCCGAGGGCGGAGATCAAGAGGTGAGGAG 1605
Db 123412 CTCACGCTGTAATCCAGCACTTTGGAGGCCGAGGGCGGAGATCAAGAGGTGAGGAG 123471
Qy 1606 ATCGAGACCATCTGGCTAAACACGGGGTGAACCCCGCTCTACTATAAAATACAAAAAT 1665
Db 123472 ATCGAGACCATCTGGCTAAACAC--GGTGAACCCCGCTCTACTATAAAATACAAAAAT 123529
Qy 1666 CAGCCGGGTGAGGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAGGCTGAGGCAGAGA 1725
Db 123530 TAGCCGGCGAGGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAGGCTGAGGCAGAGA 123589
Qy 1726 ATGGCGTGAACCCACAGGGGGCGAGCCCTACAGTGAGCCGAGATTGCGCACTGCACTCCA 1785
Db 123590 ATGGCGTGAACCCACAGGGGGCGAGCCCTGAGTGAGCCGAGATTGCGCACTGCACTCCA 123649
Qy 1786 GCCTGGAGCAGATGAGACTCGCTCTCAAAAAAAGAAAAAAGAAAAA 1838
Db 123650 GCCTGGCGCAGCAGCAGACTCGCTCTCAAAAAAAGAAAAAAGAAAAA 123702

RESULT 14
AAI99076/c
ID AAI99076 standard; DNA; 335 BP.
XX
AC AAI99076;
XX
DT 07-JAN-2002 (first entry)
XX
DE Human excretory related polynucleotide SEQ ID NO 840.

XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine;
KW excretory system; ds.
XX
OS Homo sapiens.
XX
PN WO200155313-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001323.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.

PR	14-AUG-2000;	2000US-0225213P.	PR	08-NOV-2000;	2000US-0246523P.
PR	14-AUG-2000;	2000US-0225214P.	PR	08-NOV-2000;	2000US-0246524P.
PR	14-AUG-2000;	2000US-0225266P.	PR	08-NOV-2000;	2000US-0246525P.
PR	14-AUG-2000;	2000US-0225267P.	PR	08-NOV-2000;	2000US-0246526P.
PR	14-AUG-2000;	2000US-0225268P.	PR	08-NOV-2000;	2000US-0246527P.
PR	14-AUG-2000;	2000US-0225270P.	PR	08-NOV-2000;	2000US-0246528P.
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.
PR	14-AUG-2000;	2000US-0225757P.	PR	08-NOV-2000;	2000US-0246609P.
PR	14-AUG-2000;	2000US-0225758P.	PR	08-NOV-2000;	2000US-0246610P.
PR	14-AUG-2000;	2000US-0225759P.	PR	08-NOV-2000;	2000US-0246611P.
PR	18-AUG-2000;	2000US-0226279P.	PR	08-NOV-2000;	2000US-0246613P.
PR	22-AUG-2000;	2000US-0226681P.	PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000;	2000US-0226688P.	PR	17-NOV-2000;	2000US-0249208P.
PR	22-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249214P.
PR	23-AUG-2000;	2000US-0227009P.	PR	17-NOV-2000;	2000US-0249209P.
PR	30-AUG-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249210P.
PR	01-SEP-2000;	2000US-0229287P.	PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249214P.
PR	05-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249215P.
PR	05-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249216P.
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249217P.
PR	06-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249218P.
PR	08-SEP-2000;	2000US-0231342P.	PR	17-NOV-2000;	2000US-0249244P.
PR	08-SEP-2000;	2000US-0231343P.	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	2000US-0231344P.	PR	17-NOV-2000;	2000US-0249264P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	2000US-0232080P.	PR	17-NOV-2000;	2000US-0249299P.
PR	08-SEP-2000;	2000US-0232081P.	PR	17-NOV-2000;	2000US-0249300P.
PR	12-SEP-2000;	2000US-02321968P.	PR	01-DEC-2000;	2000US-0250160P.
PR	14-SEP-2000;	2000US-02323397P.	PR	01-DEC-2000;	2000US-0250391P.
PR	14-SEP-2000;	2000US-0232398P.	PR	05-DEC-2000;	2000US-0251030P.
PR	14-SEP-2000;	2000US-0232399P.	PR	05-DEC-2000;	2000US-0251988P.
PR	14-SEP-2000;	2000US-0232400P.	PR	05-DEC-2000;	2000US-0256719P.
PR	14-SEP-2000;	2000US-0232401P.	PR	06-DEC-2000;	2000US-0251479P.
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251856P.
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251868P.
PR	14-SEP-2000;	2000US-0233065P.	PR	08-DEC-2000;	2000US-0251869P.
PR	21-SEP-2000;	2000US-0234223P.	PR	08-DEC-2000;	2000US-0251989P.
PR	21-SEP-2000;	2000US-0234274P.	PR	08-DEC-2000;	2000US-0251990P.
PR	25-SEP-2000;	2000US-0234997P.	PR	11-DEC-2000;	2000US-0254097P.
PR	25-SEP-2000;	2000US-0234998P.	PR	05-JAN-2001;	2001US-02559678P.
PR	26-SEP-2000;	2000US-0235484P.	XX		
PR	27-SEP-2000;	2000US-0235834P.	XX		
PR	27-SEP-2000;	2000US-0235836P.	XX		
PR	29-SEP-2000;	2000US-0236327P.	XX		
PR	29-SEP-2000;	2000US-0236367P.	XX		
PR	29-SEP-2000;	2000US-0236368P.	XX		
PR	29-SEP-2000;	2000US-0236369P.	XX		

```

CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 335 BP; 48 A; 107 C; 82 G; 98 T; 0 U; 0 Other;

Query Match          9.7%; Score 274.6; DB 4; Length 335;
Best Local Similarity 92.0%; Pred. No. 1.1e-60;
Matches 301; Conservative 0; Mismatches 24; Indels 2; Gaps 1;

QY 1513 TGAAGAAAAGCAGAGATTGGCGGCGACGGTGGCTCAGCGCTGTATATCCAGCATTGG 1572
Dd | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
334 TTATAAATGCGAGGAAGCGCGCGGTGGCTCAGCGCTGTATATCCAGCATTGG 275
QY 1573 GAGCCGAGGCGGCGAGATCAGAGGTCAGAGATCGAGACCATCTGGCTAAACACGGGG 1632
Dd | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
274 GAGCCGAGGCGGCTGGAATCAGAGTTCAGAGATCGAGACCATCTGGCTTAACAC--GG 217
QY 1633 TGAACCCCGTCTCTACTATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1692
Dd | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
216 TGAACCCCGTCTCTACTATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 157
QY 1693 AGTCCAGCTACTCAGGAGGCTGAGGCAAGAGATGGCGTGAACCCAGGGGGCGAGCC 1752
Dd | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
156 AGTCCAGCTACTCAGGAGGCTGAGGCAAGAGATGGCGTGAACCCAGGGGGCGAGCC 97
QY 1753 TACAGTGAGCGGAGATTGCGGCACCTGCAGCTCGACGACGACGAGTCCGCTCTC 1812
Dd | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
96 TGCAGTGAGCGGAGATTGCGGCACCTGCAGCTCGGCGAGCGAGACTCGGCTCTC 37
QY 1813 AAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1839
Dd | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
36 AAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 10

RESULT 15
AAI63426/c
ID AAI63426 standard; DNA; 335 BP.
XX
AC AAI63426;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human kidney related polynucleotide SEQ ID NO 741.
XX
KW Human; kidney antigen; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW neotropic; neuroprotective; antibacterial; virucide; fungicide;
KW ophthalmological; antiallergic; hepatotropic; antidiabetic;
KW antiinflammatory; antitumor; vulnery; anticonvulsant; antiparasitic;
KW gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; ds.
XX
OS Homo sapiens.
XX
PN WO200155323-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001343.
XX
31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR
11-JUL-2000; 2000US-0217496P.
PR
14-JUL-2000; 2000US-0218290P.
PR
26-JUL-2000; 2000US-0220963P.
PR
26-JUL-2000; 2000US-0220964P.
PR
14-AUG-2000; 2000US-0224518P.
PR
14-AUG-2000; 2000US-0224519P.
PR
14-AUG-2000; 2000US-0225213P.
PR
14-AUG-2000; 2000US-0225214P.
PR
14-AUG-2000; 2000US-0225266P.
PR
14-AUG-2000; 2000US-0225267P.
PR
14-AUG-2000; 2000US-0225268P.
PR
14-AUG-2000; 2000US-0225270P.
PR
14-AUG-2000; 2000US-0225447P.
PR
14-AUG-2000; 2000US-0225757P.
PR
14-AUG-2000; 2000US-0225758P.
PR
14-AUG-2000; 2000US-0225759P.
PR
14-AUG-2000; 2000US-0226279P.
PR
22-AUG-2000; 2000US-0226681P.
PR
22-AUG-2000; 2000US-0226686P.
PR
22-AUG-2000; 2000US-0227182P.
PR
23-AUG-2000; 2000US-0227009P.
PR
30-AUG-2000; 2000US-0228924P.
PR
01-SEP-2000; 2000US-0229287P.
PR
01-SEP-2000; 2000US-0229343P.
PR
01-SEP-2000; 2000US-0229344P.
PR
01-SEP-2000; 2000US-0229345P.
PR
05-SEP-2000; 2000US-0229509P.
PR
05-SEP-2000; 2000US-0229513P.
PR
06-SEP-2000; 2000US-0230437P.
PR
06-SEP-2000; 2000US-0230438P.
PR
08-SEP-2000; 2000US-0231242P.
PR
08-SEP-2000; 2000US-0231243P.
PR
08-SEP-2000; 2000US-0231244P.
PR
08-SEP-2000; 2000US-0231413P.
PR
08-SEP-2000; 2000US-0231414P.
PR
08-SEP-2000; 2000US-0232080P.
PR
08-SEP-2000; 2000US-0232081P.
PR
12-SEP-2000; 2000US-0231968P.
PR
14-SEP-2000; 2000US-0232397P.
PR
14-SEP-2000; 2000US-0232398P.
PR
14-SEP-2000; 2000US-0232399P.
PR
14-SEP-2000; 2000US-0232400P.
PR
14-SEP-2000; 2000US-0232401P.
PR
14-SEP-2000; 2000US-0233063P.
PR
14-SEP-2000; 2000US-0233064P.
PR
14-SEP-2000; 2000US-0233065P.
PR
21-SEP-2000; 2000US-0234223P.
PR
25-SEP-2000; 2000US-0234274P.
PR
25-SEP-2000; 2000US-0234997P.
PR
25-SEP-2000; 2000US-0234998P.
PR
26-SEP-2000; 2000US-0235484P.
PR
27-SEP-2000; 2000US-0235834P.
PR
27-SEP-2000; 2000US-0235836P.
PR
29-SEP-2000; 2000US-0236327P.
PR
29-SEP-2000; 2000US-0236367P.
PR
29-SEP-2000; 2000US-0236368P.
PR
29-SEP-2000; 2000US-0236369P.
PR
29-SEP-2000; 2000US-0236370P.
PR
02-OCT-2000; 2000US-0236802P.
PR
02-OCT-2000; 2000US-0237037P.
PR
02-OCT-2000; 2000US-0237038P.
PR
02-OCT-2000; 2000US-0237039P.
PR
02-OCT-2000; 2000US-0237040P.
PR
13-OCT-2000; 2000US-0239935P.
PR
13-OCT-2000; 2000US-0239937P.
PR
20-OCT-2000; 2000US-0240960P.
PR
20-OCT-2000; 2000US-0241221P.
PR
20-OCT-2000; 2000US-0241785P.
PR
20-OCT-2000; 2000US-0241786P.
PR
20-OCT-2000; 2000US-0241787P.
PR
20-OCT-2000; 2000US-0241808P.
PR
20-OCT-2000; 2000US-0241809P.
PR
20-OCT-2000; 2000US-0241826P.

```

PR	01-NOV-2000;	2000US-0244617P.	CC	allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
PR	08-NOV-2000;	2000US-0246474P.	CC	diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
PR	08-NOV-2000;	2000US-0246475P.	CC	arthritis and ulcerative colitis; (c) cardiovascular disorders such as
PR	08-NOV-2000;	2000US-0246476P.	CC	myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
PR	08-NOV-2000;	2000US-0246477P.	CC	cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
PR	08-NOV-2000;	2000US-0246478P.	CC	bacterial, fungal and parasitic infections. Note: The sequence data for
PR	08-NOV-2000;	2000US-0246523P.	CC	this patent did not form part of the printed specification, but was
PR	08-NOV-2000;	2000US-0246524P.	CC	obtained in electronic format directly from WIPO at
PR	08-NOV-2000;	2000US-0246525P.	CC	ftp.wipo.int/pub/published_pct_sequences
PR	08-NOV-2000;	2000US-0246526P.	XX	
PR	08-NOV-2000;	2000US-0246527P.	SQ	Sequence 335 BP; 48 A; 107 C; 82 G; 98 T; 0 U; 0 Other;
PR	08-NOV-2000;	2000US-0246528P.		
PR	08-NOV-2000;	2000US-0246532P.		
PR	08-NOV-2000;	2000US-0246609P.		
PR	08-NOV-2000;	2000US-0246610P.		
PR	08-NOV-2000;	2000US-0246611P.		
PR	08-NOV-2000;	2000US-0246613P.		
PR	17-NOV-2000;	2000US-0249207P.		
PR	17-NOV-2000;	2000US-0249208P.		
PR	17-NOV-2000;	2000US-0249209P.		
PR	17-NOV-2000;	2000US-0249210P.		
PR	17-NOV-2000;	2000US-0249211P.		
PR	17-NOV-2000;	2000US-0249212P.		
PR	17-NOV-2000;	2000US-0249213P.		
PR	17-NOV-2000;	2000US-0249214P.		
PR	17-NOV-2000;	2000US-0249215P.		
PR	17-NOV-2000;	2000US-0249216P.		
PR	17-NOV-2000;	2000US-0249217P.		
PR	17-NOV-2000;	2000US-0249218P.		
PR	17-NOV-2000;	2000US-0249219P.		
PR	17-NOV-2000;	2000US-0249244P.		
PR	17-NOV-2000;	2000US-0249245P.		
PR	17-NOV-2000;	2000US-0249264P.		
PR	17-NOV-2000;	2000US-0249265P.		
PR	17-NOV-2000;	2000US-0249297P.		
PR	17-NOV-2000;	2000US-0249298P.		
PR	17-NOV-2000;	2000US-0249300P.		
PR	01-DEC-2000;	2000US-0250160P.		
PR	01-DEC-2000;	2000US-0250391P.		
PR	05-DEC-2000;	2000US-0251030P.		
PR	05-DEC-2000;	2000US-0251988P.		
PR	05-DEC-2000;	2000US-0256719P.		
PR	06-DEC-2000;	2000US-0251479P.		
PR	08-DEC-2000;	2000US-0251856P.		
PR	08-DEC-2000;	2000US-0251868P.		
PR	08-DEC-2000;	2000US-0251869P.		
PR	08-DEC-2000;	2000US-0251989P.		
PR	08-DEC-2000;	2000US-0251990P.		
PR	11-DEC-2000;	2000US-0254097P.		
PR	05-JAN-2001;	2001US-0259678P.		
XX				
PA		(HUMA-) HUMAN GENOME SCI INC.		
PI		Rosen CA, Barash SC, Ruben SM;		
XX		WPI; 2001-488784/53.		
XX				
PT		New isolated nucleic acids and polypeptides, useful for diagnosing,		
PT		treating and/or preventing human diseases and disorders.		
XX				
PS		Disclosure; SEQ ID NO 741; 564pp + Sequence Listing; English.		
XX				
CC		The invention relates to novel kidney related polynucleotides (AAI62971-		
CC		AAI63793) and the encoded polypeptides (AAM42417-AAM42691) collectively		
CC		known as kidney antigens and the use of such kidney antigens for		
CC		detecting disorders of the kidney, especially kidney cancer and kidney		
CC		cancer metastases. The polynucleotides and proteins are also useful for		
CC		preventing, treating or ameliorating medical conditions e.g. by protein		
CC		or gene therapy. The genes are isolated from a range of human tissues		
CC		disclosed in the specification. The nucleic acids, proteins, antibodies		
CC		and (ant)agonists are useful in the diagnosis, treatment and prevention		
CC		of: (a) cancer, e.g. breast and ovarian cancer, and other cancers of the		
CC		adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,		
CC		lung, or urogenital; (b) immune disorders e.g. Addison's disease,		

Query Match	9.7%;	Score 274.6;	DB 5;	Length 335;
Best Local Similarity	92.0%;	Pred. No. 1.1e-60;		
Matches 301;	Conservative	0;	Mismatches 24;	Indels 2; Gaps 1;
QY	1513	TGAAGAAAAGCAGGAGTTGCGGGCACGGTGGCTCACGCCCTGTAATCCACACATTTGG	1572	
DB	334	TTATAAATGCGAGGAAGCGCGCGGTGGCTCACGCCCTGTAATCCACACATTTGG	275	
QY	1573	GAGGCCGAGCGGGCAGATCACGAGTTCAGGAGATCGAGACCATCTCTGCTAACACGGGG	1632	
DB	274	GAGGCCGAGCGGGTGGATCACGAGTTCAGGAGATCGAGACCATCTCTGCTAACAC--GG	217	
QY	1633	TGAAACCCCGTCTTACTATAAAATACAAAAATACAAAAATACAAAAATACAAAAAT	1692	
DB	216	TGAAACCCCGTCTTACTATAAAATACAAAAATACAAAAATACAAAAATACAAAAAT	157	
QY	1693	AGTCCCGCTACTCAGGAGGCTGAGGCAAGAGATGGCGTGAACCCAGGGGGCGGAGCC	1752	
DB	156	AGTCCCGCTACTCAGGAGGCTGAGGCAAGAGATGGCGTGAACCCAGGGGGCGGAGCC	97	
QY	1753	TACAGTGAGCGGAGATTGCGGCACCTGCACCTCCAGCTCGACACAGTGCAGACTCCGTCTC	1812	
DB	96	TGCAGTGAGCGGAGATTGCGGCACCTGCACCTCCAGCTCGGCGACAGAGACTCCGTCTC	37	
QY	1813	AAAAAAG	1839	
DB	36	AAAAAAG	10	

Search completed: June 28, 2005, 19:33:53
Job time : 994.126 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 18:22:52 ; Search time 320.226 Seconds
(without alignments)
14414.621 Million cell updates/sec

Title: US-08-731-499-6

Perfect score: 2821

Sequence: 1 ATCTAAGACGACAGCCTG.....AGGGTTCAAGACGACGCTG 2821

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2821	100.0	2821	2	US-08-680-395-6
2	2821	100.0	2821	4	US-08-892-693-6
3	270.8	9.6	601	4	US-09-949-016-174743
4	270.6	9.6	255679	4	US-09-949-016-17189
5	269.6	9.6	51508	4	US-09-949-016-16681
6	269.4	9.5	131332	4	US-09-949-016-15535
7	267.4	9.5	258775	4	US-09-949-016-15435
8	267.2	9.5	24166	4	US-09-949-016-13755
9	267.2	9.5	91665	4	US-09-949-016-12334
10	266.6	9.5	44166	4	US-09-949-016-15829
11	266.6	9.5	54180	4	US-09-949-016-14894
12	266.4	9.4	67479	4	US-09-949-016-11804
13	266.4	9.4	71119	4	US-09-949-016-15358
14	263.6	9.3	86273	4	US-09-949-016-15273
15	263	9.3	360470	4	US-09-949-016-13173
16	262.8	9.3	18048	4	US-09-949-016-12363
17	262.8	9.3	18049	4	US-09-949-016-12635
18	262.8	9.3	136917	4	US-09-949-016-16369
19	262.4	9.3	601	4	US-09-949-016-23990
20	262.4	9.3	601	4	US-09-949-016-160457
21	259.8	9.2	52496	4	US-09-949-016-16118
22	259.8	9.2	52496	4	US-09-949-016-16119
23	259.6	9.2	462589	4	US-09-949-016-16119
24	259.6	9.2	476044	4	US-09-949-016-12900
25	259.2	9.2	601	4	US-09-949-016-12412
26	257.4	9.1	601	4	US-09-949-016-40745
27	256.2	9.1	15725	4	US-09-949-016-91528
					Sequence 17607, A

c 28	256.2	9.1	106199	4	US-09-949-016-12393	Sequence 12393, A
c 29	256.2	9.1	107980	4	US-09-949-016-14370	Sequence 14370, A
c 30	256	9.1	77100	4	US-09-949-016-16418	Sequence 16418, A
c 31	256	9.1	87863	4	US-09-949-016-14402	Sequence 14402, A
c 32	255.8	9.1	601	4	US-09-949-016-91529	Sequence 91529, A
c 33	255	9.0	84558	4	US-09-949-016-15752	Sequence 15752, A
c 34	254.6	9.0	19091	4	US-09-949-016-15805	Sequence 15805, A
c 35	254.2	9.0	601	4	US-09-949-016-143877	Sequence 143877, A
c 36	253.6	9.0	78846	4	US-09-949-016-12396	Sequence 12396, A
c 37	253.6	9.0	78846	4	US-09-949-016-12791	Sequence 12791, A
c 38	253.6	9.0	78846	4	US-09-949-016-12792	Sequence 12792, A
c 39	253.6	9.0	78846	4	US-09-949-016-12793	Sequence 12793, A
c 40	253.6	9.0	78850	4	US-09-949-016-16013	Sequence 16013, A
c 41	253.6	9.0	78850	4	US-09-949-016-16014	Sequence 16014, A
c 42	253.6	9.0	78850	4	US-09-949-016-16015	Sequence 16015, A
c 43	253.6	9.0	78850	4	US-09-949-016-16016	Sequence 16016, A
c 44	253.6	9.0	78850	4	US-09-949-016-16201	Sequence 16201, A
c 45	253.6	9.0	78850	4	US-09-949-016-16202	Sequence 16202, A

ALIGNMENTS

RESULT 1
US-08-680-395-6
; Sequence 6, Application US/08680395
; Patent No. 5892010
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W.
; APPLICANT: Collins, Colin
; APPLICANT: Hwang, Soo-in
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowbel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: Genes from the 20q13 Amplicon and Their
; TITLE OF INVENTION: Uses
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/680,395
; FILING DATE: 15-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-06890005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2821 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..2821
; OTHER INFORMATION: /note= "cdna clone GCAP encodes a
; OTHER INFORMATION: guanine cyclase activating protein
; OTHER INFORMATION:

US-08-680-395-6									
Query Match 100.0%; Score 2821; DB 2; Length 2821;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 2821; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	ATCCTTAAGACGCACAGCCTGGGAAGCCAGCACTGGGGAAGTGGTGCTCAGGGGATGTGGGT	60						
Db	1	ATCCTTAAGACGCACAGCCTGGGAAGCCAGCACTGGGGAAGTGGTGCTCAGGGGATGTGGGT	60						
Qy	61	CAC TGGGGTGAAGTGGAGCTTTTCAGGGTCTCCCGTCAATGACGTGATGTTTCTTTGGC	120						
Db	61	CAC TGGGGTGAAGTGGAGCTTTTCAGGGTCTCCCGTCAATGACGTGATGTTTCTTTGGC	120						
Qy	121	AGGGAATTTTACAGCTGAAGAAAGCCTGCCGCGAGAGCTACAACTGACGCAAGGCCAGC	180						
Db	121	AGGGAATTTTACAGCTGAAGAAAGCCTGCCGCGAGAGCTACAACTGACGCAAGGCCAGC	180						
Qy	181	TGCTCACACCCGAGGAGTCTGTGACAGAGTCTTCTCCTCGTGGATGAGAAATGGAGATG	240						
Db	181	TGCTCACACCCGAGGAGTCTGTGACAGAGTCTTCTCCTCGTGGATGAGAAATGGAGATG	240						
Qy	241	GTAAGAGGGCAGAGATGGGAGAGTGTGTCCACTCTGCATCATCGCCACTTTCTGGCC	300						
Db	241	GTAAGAGGGCAGAGATGGGAGAGTGTGTCCACTCTGCATCATCGCCACTTTCTGGCC	300						
Qy	301	GCAGTCTCTGGGCAAGGCCCTCCACCTTCCAACTCCCTGGGGTCTCATCTCTGAGAGGC	360						
Db	301	GCAGTCTCTGGGCAAGGCCCTCCACCTTCCAACTCCCTGGGGTCTCATCTCTGAGAGGC	360						
Qy	361	TGTGAGAAAGATGTCAATGAACTTAACTAAAGGAGTCTCATGAGCACGTGTTGTAGGAGTGAC	420						
Db	361	TGTGAGAAAGATGTCAATGAACTTAACTAAAGGAGTCTCATGAGCACGTGTTGTAGGAGTGAC	420						
Qy	421	TAAAGTCCTTACAGAGTTGCTGATGGAGGCGAGGCAAGGCAAGATAGAGAAATAGGAAAC	480						
Db	421	TAAAGTCCTTACAGAGTTGCTGATGGAGGCGAGGCAAGGCAAGATAGAGAAATAGGAAAC	480						
Qy	481	TTTGGAGTCAGGAGGAGTGCATATTTAGCTTCTCGTCTAGTCTCAATTTCTCTCATC	540						
Db	481	TTTGGAGTCAGGAGGAGTGCATATTTAGCTTCTCGTCTAGTCTCAATTTCTCTCATC	540						
Qy	541	TGGAAATGGGGTAATTAATAGTGGTTGAGAGGAATGAATAGGATAATGTGTTTAAAGAGC	600						
Db	541	TGGAAATGGGGTAATTAATAGTGGTTGAGAGGAATGAATAGGATAATGTGTTTAAAGAGC	600						
Qy	601	AGGCATAGGGTAGACCTCCATTCAGGCTGTCTGGGCTTTCTCCCTGTAGCCCAAGGCC	660						
Db	601	AGGCATAGGGTAGACCTCCATTCAGGCTGTCTGGGCTTTCTCCCTGTAGCCCAAGGCC	660						
Qy	661	AGCCTCAGGGCTATGTGGGAGAGAGCTGGCTTGGAAATACACACTTGAGCCCTCCAGCTC	720						
Db	661	AGCCTCAGGGCTATGTGGGAGAGAGCTGGCTTGGAAATACACACTTGAGCCCTCCAGCTC	720						
Qy	721	TCTCAGCTCCACCCAGCATTTCCGTGGTACCATGGCAAAAGTAAAACTTCAATTCATCA	780						
Db	721	TCTCAGCTCCACCCAGCATTTCCGTGGTACCATGGCAAAAGTAAAACTTCAATTCATCA	780						
Qy	781	GCAAGAAAGCCCTTAAAGGTGGCAGAGACTCTCTGGAGATTCAGACCTGTGCAAGCC	840						
Db	781	GCAAGAAAGCCCTTAAAGGTGGCAGAGACTCTCTGGAGATTCAGACCTGTGCAAGCC	840						
Qy	841	GCAAGCTTGAGGTCTGAGACTGCAGGATGTTGGCATGAAGACGTGTAGCGCATCTCTGGG	900						
Db	841	GCAAGCTTGAGGTCTGAGACTGCAGGATGTTGGCATGAAGACGTGTAGCGCATCTCTGGG	900						
Qy	901	AGCAGAGTCTCTCTCTCTGCCCCCAGACCCAGGCTCTCCCTTCTTCTACATGACCACTC	960						
Db	901	AGCAGAGTCTCTCTCTCTGCCCCCAGACCCAGGCTCTCCCTTCTTCTACATGACCACTC	960						
Qy	961	TCCTCCCCCTTGTCTCAGGCGAGCTGTCTGAACTGAGTTTGTGAAAGTGTCCCGTCGGGA	1020						
Db	961	TCCTCCCCCTTGTCTCAGGCGAGCTGTCTGAACTGAGTTTGTGAAAGTGTCCCGTCGGGA	1020						

Qy	1021	CAAGTGGGTGATGAAGATGCTGCAGATGGAATCCCCAGCAGCTGCCTCGCTCAGCA	1080						
Db	1021	CAAGTGGGTGATGAAGATGCTGCAGATGGAATCCCCAGCAGCTGCCTCGCTCAGCA	1080						
Qy	1081	GAGACGGAAAAAGTGCATGTTTCTGAGGAGTCTGGGGCCCCCTCCACGACTCCAGGGCTACC	1140						
Db	1081	GAGACGGAAAAAGTGCATGTTTCTGAGGAGTCTGGGGCCCCCTCCACGACTCCAGGGCTACC	1140						
Qy	1141	CAGTTCCTCAGGCTAGTAGGAGGTCCTCTGGCTCAGCTGCTCATGCCCCACTCTTCCCC	1200						
Db	1141	CAGTTCCTCAGGCTAGTAGGAGGTCCTCTGGCTCAGCTGCTCATGCCCCACTCTTCCCC	1200						
Qy	1201	TGTTGTTGACTTCTTGGCACCCTCTGTGCAGGCTGTAGTGGGATGCGGAAGGGCTGCTG	1260						
Db	1201	TGTTGTTGACTTCTTGGCACCCTCTGTGCAGGCTGTAGTGGGATGCGGAAGGGCTGCTG	1260						
Qy	1261	GGTTTGAAGTGGCCCAACAGGGCATAGTCCATTTTGGAGGAGTCCCTGGGATGGTGAAGGG	1320						
Db	1261	GGTTTGAAGTGGCCCAACAGGGCATAGTCCATTTTGGAGGAGTCCCTGGGATGGTGAAGGG	1320						
Qy	1321	AATTCAGTTACTTTTCTGTTTCTGTTTCTGGGAGTCTGTGCTTTGGCTGGGTGGT	1380						
Db	1321	AATTCAGTTACTTTTCTGTTTCTGTTTCTGGGAGTCTGTGCTTTGGCTGGGTGGT	1380						
Qy	1381	GTGGGCTCCACACAGTCTTCTGGGTGTTCTCAGTTTGGAAAGCAAGAGCCAACTGAGGGGTGA	1440						
Db	1381	GTGGGCTCCACACAGTCTTCTGGGTGTTCTCAGTTTGGAAAGCAAGAGCCAACTGAGGGGTGA	1440						
Qy	1441	GGGTCCCAACAGACCAAAATCAGAAATAGAAACACAAAGACTCGTAGGAGGCGGGGTGGGA	1500						
Db	1441	GGGTCCCAACAGACCAAAATCAGAAATAGAAACACAAAGACTCGTAGGAGGCGGGGTGGGA	1500						
Qy	1501	GGGTGTTGAGACTGAGAAAGGCGAGGTTTCCGGGCGAGCTGCTCAGCCCTGTAATC	1560						
Db	1501	GGGTGTTGAGACTGAGAAAGGCGAGGTTTCCGGGCGAGCTGCTCAGCCCTGTAATC	1560						
Qy	1561	CCAGCCTTTTGGGAGGCGAGGCGGCGAGATCAAGAGTTCAGGAGTCCAGACCATCTCTG	1620						
Db	1561	CCAGCCTTTTGGGAGGCGAGGCGGCGAGATCAAGAGTTCAGGAGTCCAGACCATCTCTG	1620						
Qy	1621	GCTAACCGGGGTGAAACCCCGTCTCTACTAAAAATACAAAAATCAGCCGGGTGAGGTG	1680						
Db	1621	GCTAACCGGGGTGAAACCCCGTCTCTACTAAAAATACAAAAATCAGCCGGGTGAGGTG	1680						
Qy	1681	CGGGCGGCTGTAGTCCAGCTACTCAGGAGGCTGAGGCAAGAGAAATGGCGTGAACCCCA	1740						
Db	1681	CGGGCGGCTGTAGTCCAGCTACTCAGGAGGCTGAGGCAAGAGAAATGGCGTGAACCCCA	1740						
Qy	1741	GGGGCGGAGCTTACAGTGAGCCGAGATTGGCCACTGCACTCCAGCCTGGACGACAGTG	1800						
Db	1741	GGGGCGGAGCTTACAGTGAGCCGAGATTGGCCACTGCACTCCAGCCTGGACGACAGTG	1800						
Qy	1801	AGACTCCGTCTCAAAAAAAGAAAAAGAAAGGAGGAGTGGGGGGGAG	1860						
Db	1801	AGACTCCGTCTCAAAAAAAGAAAAAGAAAGGAGGAGTGGGGGGGAG	1860						
Qy	1861	GGGGCAGCAATAATTTCTATACTTCCGGATGCTGAGGGGCGTTTCATGGGGAGGACCCCTG	1920						
Db	1861	GGGGCAGCAATAATTTCTATACTTCCGGATGCTGAGGGGCGTTTCATGGGGAGGACCCCTG	1920						
Qy	1921	GCCTCTCTCTCCCAAGGCATCTCACAGTGGTGTCAAAGGAAAAAATGGCAGCAATAA	1980						
Db	1921	GCCTCTCTCTCCCAAGGCATCTCACAGTGGTGTCAAAGGAAAAAATGGCAGCAATAA	1980						
Qy	1981	CGCTCAGGCTGTGGTCTTTCTGCTTTGAAAGGGTACGTGTACTTAAAGGAGCTGTTT	2040						
Db	1981	CGCTCAGGCTGTGGTCTTTCTGCTTTGAAAGGGTACGTGTACTTAAAGGAGCTGTTT	2040						
Qy	2041	CAGCTCTGCTGGGTGCTGCTGCGACCCCTCTGCGCAACCCCACTCCCCCAACA	2100						
Db	2041	CAGCTCTGCTGGGTGCTGCTGCGACCCCTCTGCGCAACCCCACTCCCCCAACA	2100						

Qy	2101	TCCTCTCTTTCCATCCGATATCCCCAGTAGTGAGCCTTCCAACTCCAGGCCATAAGCTG	2160
Db	2101	TCCTCTCTTTCCATCCGATATCCCCAGTAGTGAGCCTTCCAACTCCAGGCCATAAGCTG	2160
Qy	2161	AATGTTTCTCTTTAAAGGATGGAGAAACTTCTCTGTGTCTCTGGCAAGAAATTTGGGGGAC	2220
Db	2161	AATGTTTCTCTTTAAAGGATGGAGAAACTTCTCTGTGTCTCTGGCAAGAAATTTGGGGGAC	2220
Qy	2221	TGTTGACTGGGATTTGTGGGCTTGGGCTTGGCTTTAACTGCTGTGTGTGACCACCAAGCAGCCCA	2280
Db	2221	TGTTGACTGGGATTTGTGGGCTTGGGCTTGGCTTTAACTGCTGTGTGTGACCACCAAGCAGCCCA	2280
Qy	2281	CTTCTCCTCCCTAACCTTGGTTATGTCTTGGCAGCACAGTGAGCAGGTCCGACTAGGCGCA	2340
Db	2281	CTTCTCCTCCCTAACCTTGGTTATGTCTTGGCAGCACAGTGAGCAGGTCCGACTAGGCGCA	2340
Qy	2341	ACAGTTTTGGATTAATGTGTTTTTGTAGATGTGGAATATTTTTTTGTATATAAACTCTTAT	2400
Db	2341	ACAGTTTTGGATTAATGTGTTTTTGTAGATGTGGAATATTTTTTTGTATATAAACTCTTAT	2400
Qy	2401	GTGTAAACCCCAATATAGAAACTAGATTAAAGGGAGTCTCTCTGGTTGAAAGGGGAGCTG	2460
Db	2401	GTGTAAACCCCAATATAGAAACTAGATTAAAGGGAGTCTCTCTGGTTGAAAGGGGAGCTG	2460
Qy	2461	AGTACCCTCTGGAACTGGAGCCACTCTGAAAAAAGCAAACTGAAAAACAGTGCCCTGGG	2520
Db	2461	AGTACCCTCTGGAACTGGAGCCACTCTGAAAAAAGCAAACTGAAAAACAGTGCCCTGGG	2520
Qy	2521	TCACGTGTTACTCCCTATAAGACAGTTTAAAGTGAGACCTCGAAAAACAACTTTTACCTT	2580
Db	2521	TCACGTGTTACTCCCTATAAGACAGTTTAAAGTGAGACCTCGAAAAACAACTTTTACCTT	2580
Qy	2581	GAATAGATAGGTTTTTATGTTTGGTATATAAGAAATAAAACTTAACCTATTAAACCTCGAGAC	2640
Db	2581	GAATAGATAGGTTTTTATGTTTGGTATATAAGAAATAAAACTTAACCTATTAAACCTCGAGAC	2640
Qy	2641	TTTACAGGTGTGTTATTTTCATATGATAGTCAATATAAAATTTCCCTTTAGACATCAATTTTA	2700
Db	2641	TTTACAGGTGTGTTATTTTCATATGATAGTCAATATAAAATTTCCCTTTAGACATCAATTTTA	2700
Qy	2701	GGTAAAAATAATTGATTAGAAAAATATTGGCCAGGTGACAGCTCACACCTGCAATCC	2760
Db	2701	GGTAAAAATAATTGATTAGAAAAATATTGGCCAGGTGACAGCTCACACCTGCAATCC	2760
Qy	2761	CAGGACTTTTGGAGCCGAGCGGGTGGATCACTGTAGGTTCAGGGGTTCAAGACCAGCCT	2820
Db	2761	CAGGACTTTTGGAGCCGAGCGGGTGGATCACTGTAGGTTCAGGGGTTCAAGACCAGCCT	2820
Qy	2821	G 2821	
Db	2821	G 2821	

RESULT 2

US-08-892-695-6

03-08-892-093-8
; Sequence 6, Application US/08892695A

; Patent No. 6808878

GENERAL INFORMATION:

APPLICANT: Gray, Joe W

; APPLICANT: Collins, Collin

APPLICANT: Hwang, Soo In.

APPLICANT: Godfrey, Tony

APPLICANT: Kowel, David

; APPLICANT: Rommens, Johanna

;
; TITLE OF INVENTION: GENES FROM THE 20Q13 AMPLICON AND THEIR USES

; FILE REFERENCE: 2500.124US3

; CURRENT APPLICATION NUMBER: US/08/892,695A

; CURRENT FILING DATE: 1997-07-15

; EARLIER APPLICATION NUMBER: 08/785,532
 ; EARLIER FILING DATE: 1997-01-17

; EARLIER FILING DATE: 1997-01-17
 ; EARLIER APPLICATION NUMBER: 08/721 488

EARLIER APPLICATION NUMBER: 08/731,499
EARLIER FILING DATE: 1996-10-16

; EARLIER FILING DATE: 1996-10-16
 ; EARLIER APPLICATION NUMBER: 08/680 395

```

; EARLIER FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 2821
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:GCAP
US-08-892-695-6

```

Query Match	100.0%;	Score 2821;	DB 4;	Length 2821;
Best Local Similarity	100.0%;	Prod. No. 0;		
Matches 2821;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATCCTAAGACGCACAGCCTCGGAAGCCAGACATGGGGAAAGTGGTGTCTGAGGGAATGTGGGT	60	
Db	1	ATCCTAAGACGCACAGCCTCGGAAGCCAGACATGGGGAAAGTGGTGTCTGAGGGAATGTGGGT	60	
Qy	61	CATCGGGTGAAGGTGAGCTTTTCAGGGTCTCCGGTCAATGCAGCTGAGTTTTCTTTGGC	120	
Db	61	CATCGGGTGAAGGTGAGCTTTTCAGGGTCTCCGGTCAATGCAGCTGAGTTTTCTTTGGC	120	
Qy	121	AGGGAATTTACCAAGCTCAAGAAAGCCCTCCGGCGAGAGCTACAACCTGAGCAAGGCCAGC	180	
Db	121	AGGGAATTTACCAAGCTCAAGAAAGCCCTCCGGCGAGAGCTACAACCTGAGCAAGGCCAGC	180	
Qy	181	TGCTCACACCCGAGGAGGTGGTGACAGGATCTTCTCCTGGTGGATGAGAATGGAGATG	240	
Db	181	TGCTCACACCCGAGGAGGTGGTGACAGGATCTTCTCCTGGTGGATGAGAATGGAGATG	240	
Qy	241	GTAAGAGGGGAGAGATGGGGAGAGTCTGTCTCACTCTGCATCATCGCCACTTCTCGCC	300	
Db	241	GTAAGAGGGGAGAGATGGGGAGAGTCTGTCTCACTCTGCATCATCGCCACTTCTCGCC	300	
Qy	301	GCAGCTCTTTGGGCAAGGCCCTCCACCTTCCAACCTGGGGTCTCATCTGTGAGAAAGGC	360	
Db	301	GCAGCTCTTTGGGCAAGGCCCTCCACCTTCCAACCTGGGGTCTCATCTGTGAGAAAGGC	360	
Qy	361	TGTGGAGAAGATGTCATGAATCAACAAAGGAGCTCATGAGCACGTGTTGTAGAGTGAC	420	
Db	361	TGTGGAGAAGATGTCATGAATCAACAAAGGAGCTCATGAGCACGTGTTGTAGAGTGAC	420	
Qy	421	TAAAGTCTTACAGGAGTTCGTGATGAGGCCAGGCAGCAGAGATGAAGAATAGGAC	480	
Db	421	TAAAGTCTTACAGGAGTTCGTGATGAGGCCAGGCAGCAGAGATGAAGAATAGGAC	480	
Qy	481	TTTGGAGTCAGGAGGAGTGATATATTGAGCTTCTCGTCTAGTCTCAATTTCTCTCATC	540	
Db	481	TTTGGAGTCAGGAGGAGTGATATATTGAGCTTCTCGTCTAGTCTCAATTTCTCTCATC	540	
Qy	541	TGGAATAATGGGATATAATAGTGGTTGAGAGGAATGAATAGGATAATGTGTTTAAGAGC	600	
Db	541	TGGAATAATGGGATATAATAGTGGTTGAGAGGAATGAATAGGATAATGTGTTTAAGAGC	600	
Qy	601	AGGCATAGGCTAGACCTTCAATCAGGCTGCTTGGGCTTCTCCTCGTCTGAGCCAAAGCC	660	
Db	601	AGGCATAGGCTAGACCTTCAATCAGGCTGCTTGGGCTTCTCCTCGTCTGAGCCAAAGCC	660	
Qy	661	AGCCTCAGGGCTATGTGGGAGAGAGCTGGCTTGGAAATACACATTGAGCCCTCCAGCTC	720	
Db	661	AGCCTCAGGGCTATGTGGGAGAGAGCTGGCTTGGAAATACACATTGAGCCCTCCAGCTC	720	
Qy	721	TCTCAGCTCCACCAGCATTTCCGTGTGATCCATCGCGCAAAAGTAAACTTCATTCATCA	780	
Db	721	TCTCAGCTCCACCAGCATTTCCGTGTGATCCATCGCGCAAAAGTAAACTTCATTCATCA	780	
Qy	781	GCAAGAAAGCCCTTAAAGGTGCGAGAGACTCTCGAGATTCAGACACTCTGCAAGACC	840	
Db	781	GCAAGAAAGCCCTTAAAGGTGCGAGAGACTCTCGAGATTCAGACACTCTGCAAGACC	840	
Qy	841	GCAAGCTTGAGGCTTGAGACTGCAAGGATAGTTGGCATAGAAGTGTAGGCGCATCTCTGGG	900	

Db 841 GCAAGCTTGAGGTCTGAGACTGACGATAGTTGGCATAGACGCTGTAGGCGCATCCTGGG 900
 QY 901 AGCAGAGTCTCTCTCTGCCCCCAGACCCAGGCTCTCCCTTCTTCTACATGACCACTC 960
 Db 901 AGCAGAGTCTCTCTCTGCCCCCAGACCCAGGCTCTCCCTTCTTCTACATGACCACTC 960
 QY 961 TCCTCCCTCTGCTCAGGCGAGTCTCTCTGAAAGAGTTGTTGAAGGTGCGCGTCGGGA 1020
 Db 961 TCCTCCCTCTGCTCAGGCGAGTCTCTCTGAAAGAGTTGTTGAAGGTGCGCGTCGGGA 1020
 QY 1021 CAAGTGGGTGATGAAGTCTGTCAGATGGAATCCAGCAGCTGGCTCGCTCAGCA 1080
 Db 1021 CAAGTGGGTGATGAAGTCTGTCAGATGGAATCCAGCAGCTGGCTCGCTCAGCA 1080
 QY 1081 GAGACGGAAGATGCCATGTTCTGAGGAGTCTGGGGCCCCCTCCAGCAGCTCCAGGCTCACC 1140
 Db 1081 GAGACGGAAGATGCCATGTTCTGAGGAGTCTGGGGCCCCCTCCAGCAGCTCCAGGCTCACC 1140
 QY 1141 CAGGTTTCCAGGGTAGTAGGAGGTCCCTCGCTCAGGCTGCTCATGCCCACTCTTCCCC 1200
 Db 1141 CAGGTTTCCAGGGTAGTAGGAGGTCCCTCGCTCAGGCTGCTCATGCCCACTCTTCCCC 1200
 QY 1201 TGGTGTGACTTCTTCTGCAACCCCTGTCAGGCTGAGTGGGATGGGAAAGGCTGCTG 1260
 Db 1201 TGGTGTGACTTCTTCTGCAACCCCTGTCAGGCTGAGTGGGATGGGAAAGGCTGCTG 1260
 QY 1261 GGTGTGAGTGGCCNACAGGCGTAGTCCATTTTGGAGGAGTCCCTGGGATGGTGAAGG 1320
 Db 1261 GGTGTGAGTGGCCNACAGGCGTAGTCCATTTTGGAGGAGTCCCTGGGATGGTGAAGG 1320
 QY 1321 AATTCAGTTACTTTTCTGTTTCCAGCGCTCTCGGAGGAGTGTGCTTGGCTGGGTGGTT 1380
 Db 1321 AATTCAGTTACTTTTCTGTTTCCAGCGCTCTCGGAGGAGTGTGCTTGGCTGGGTGGTT 1380
 QY 1381 GTGGGCTCCACAGTTTCTGGGTCTCTCAGTTTGAAGAGCAAGACCACTGAGGGGTGA 1440
 Db 1381 GTGGGCTCCACAGTTTCTGGGTCTCTCAGTTTGAAGAGCAAGACCACTGAGGGGTGA 1440
 QY 1441 GGGTCCCAACAGCCNATCAGAAATGAGAAACAAAGACTGGTAGGAGGCGAGGGTGGGA 1500
 Db 1441 GGGTCCCAACAGCCNATCAGAAATGAGAAACAAAGACTGGTAGGAGGCGAGGGTGGGA 1500
 QY 1501 GGGTGTGAGACTGAGAAAGGCGAGGAGTTGCGGGCAGCGTGGCTCAGCCCTGTAATC 1560
 Db 1501 GGGTGTGAGACTGAGAAAGGCGAGGAGTTGCGGGCAGCGTGGCTCAGCCCTGTAATC 1560
 QY 1561 CCAGCACTTTGGAGGCGGAGCGGCGAGATCAGAGGTCAGGAGATCGAGACCATCCTG 1620
 Db 1561 CCAGCACTTTGGAGGCGGAGCGGCGAGATCAGAGGTCAGGAGATCGAGACCATCCTG 1620
 QY 1621 GCTAACAGGGGTGAAACCCGCTCTCTACTAAATACAAATATCAGCCGGGTGAGGTG 1680
 Db 1621 GCTAACAGGGGTGAAACCCGCTCTCTACTAAATACAAATATCAGCCGGGTGAGGTG 1680
 QY 1681 GCGGCGGCTGTAGTCCAGCTACTCAGAGGCTGAGGCAAGAGAAATGCGCTGAACCCCA 1740
 Db 1681 GCGGCGGCTGTAGTCCAGCTACTCAGAGGCTGAGGCAAGAGAAATGCGCTGAACCCCA 1740
 QY 1741 GGGGCGGAGCTCAGTGAGCCGAGATTGCGCACTGCACTCCAGCTCGACAGCAGTGTG 1800
 Db 1741 GGGGCGGAGCTCAGTGAGCCGAGATTGCGCACTGCACTCCAGCTCGACAGCAGTGTG 1800
 QY 1801 AGACTCCCTCTCAAAAAAAGAAAGAAAGAAAGGAGGAGGAGTTTGGGGGCGAG 1860
 Db 1801 AGACTCCCTCTCAAAAAAAGAAAGAAAGAAAGGAGGAGGAGTTTGGGGGCGAG 1860
 QY 1861 GGGGCGAGCAATAATCTATACTCCGGGAGTCTGAGGGGCTTCATCGGAGGAGGAGGCTG 1920
 Db 1861 GGGGCGAGCAATAATCTATACTCCGGGAGTCTGAGGGGCTTCATCGGAGGAGGAGGCTG 1920
 QY 1921 GCCTCTCTCTCCCAAGGAGTCTCTCAGGAGTGTCAACAGGAAATGCGAGCAATA 1980
 Db 1921 GCCTCTCTCTCCCAAGGAGTCTCTCAGGAGTGTCAACAGGAAATGCGAGCAATA 1980

QY 1981 CGCTGCAGGCTGTGCTCTTCTGCTCTTGAAGGCTCAGCTGTACTTAAAGGAGCTGTTT 2040
 Db 1981 CGCTGCAGGCTGTGCTCTTCTGCTCTTGAAGGCTCAGCTGTACTTAAAGGAGCTGTTT 2040
 QY 2041 CAGCTCTGCTGCTGCTGCTCTGCGGACCCCTCTGTCACCAACCCACCACTCCCCCAACA 2100
 Db 2041 CAGCTCTGCTGCTGCTGCTCTGCGGACCCCTCTGTCACCAACCCACCACTCCCCCAACA 2100
 QY 2101 TCCTCTCTTTCATCATATCCCCCAGTATGGAACCTTCCACAACTCCAGCAGCATAGCTG 2160
 Db 2101 TCCTCTCTTTCATCATATCCCCCAGTATGGAACCTTCCACAACTCCAGCAGCATAGCTG 2160
 QY 2161 AATGTTTCTCTTTAAAGGATGGAAGAACTTCTGCTGCTCTGCAAGAAATGGGGGAC 2220
 Db 2161 AATGTTTCTCTTTAAAGGATGGAAGAACTTCTGCTGCTCTGCAAGAAATGGGGGAC 2220
 QY 2221 TGTGACTGGGATTTGCTGGCTGGGCTTGAACCTTCTAAGCTGTGTGACCCCAAGACAGCA 2280
 Db 2221 TGTGACTGGGATTTGCTGGCTGGGCTTGAACCTTCTAAGCTGTGTGACCCCAAGACAGCA 2280
 QY 2281 CTTCTCTCCCTAACTTGTATGCTTGGCAGCAGTGTGCGAGCTCGGACTAGGCGA 2340
 Db 2281 CTTCTCTCCCTAACTTGTATGCTTGGCAGCAGTGTGCGAGCTCGGACTAGGCGA 2340
 QY 2341 ACAGTTTTGGATTAATGTTGTTTTTAGATGTGGAATTTATTTTGTATATAAATCTTTAT 2400
 Db 2341 ACAGTTTTGGATTAATGTTGTTTTTAGATGTGGAATTTATTTTGTATATAAATCTTTAT 2400
 QY 2401 GTGTAACCCCAATATAGAACTAGATTAAAGGGAGTCTCTCTGTTGAAAGGGGAGCTG 2460
 Db 2401 GTGTAACCCCAATATAGAACTAGATTAAAGGGAGTCTCTCTGTTGAAAGGGGAGCTG 2460
 QY 2461 AGTACCTCTCGAACTGAGGAGCCTCTGAAAGAAAGCAAACTGAAACCACTGCGCTGG 2520
 Db 2461 AGTACCTCTCGAACTGAGGAGCCTCTGAAAGAAAGCAAACTGAAACCACTGCGCTGG 2520
 QY 2521 TCAGTGTACTCTCTATAGACAGTTTAAAGTGAGACCTGGAAGAAACATTTGCTTTACCTT 2580
 Db 2521 TCAGTGTACTCTCTATAGACAGTTTAAAGTGAGACCTGGAAGAAACATTTGCTTTACCTT 2580
 QY 2581 GAATAGATAGTGTGTTTATGTTGTTATATAGAAATATAAATCTAATTAACCTTGAGAC 2640
 Db 2581 GAATAGATAGTGTGTTTATGTTGTTATATAGAAATATAAATCTAATTAACCTTGAGAC 2640
 QY 2641 TTTACAGTGTGTTTATGTTTATATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
 Db 2641 TTTACAGTGTGTTTATGTTTATATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
 QY 2701 GGTAAAAAATAATGATTAGAAAAATATTTGGCCAGGTGAGAGCTCACACTGCAATCC 2760
 Db 2701 GGTAAAAAATAATGATTAGAAAAATATTTGGCCAGGTGAGAGCTCACACTGCAATCC 2760
 QY 2761 CAGGACTTTGGAGCGGAGCGGCTGGATCCTCTGAGGTGAGGGGTTCAAGACAGGCT 2820
 Db 2761 CAGGACTTTGGAGCGGAGCGGCTGGATCCTCTGAGGTGAGGGGTTCAAGACAGGCT 2820
 QY 2821 G 2821
 Db 2821 G 2821

RESULT 3

US-09-949-016-174743/c
 ; Sequence 174743, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CLO01307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14

```

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 174743
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-174743

Query Match
Best Local Similarity 9.6%; Score 270.8; DB 4; Length 601;
Matches 299; Conservative 1; Mismatches 28; Indels 2; Gaps 1;

QY 1509 AGACTGAAGAAAGCAGGAGTTGGCGGCACCGGTGGCTCAGCCTGTAAATCCGAGCACT 1568
DB 465 ATAATAATAATAAAAAAACAAGCGCGCGCGGTGGCTCAGCCTGTAAATCCGAGCACT 406
QY 1569 TTGGGAGGCGGAGGCGGCGAGATCAGAGGTGAGGAGATCGAGCATCTCTGGCTAAACAC 1628
DB 405 TTGGGAGGCGGAGGCGGCGGATCAGAGGTGAGGAGATCAAGACATCTCTGGCTAAACAC 346
QY 1629 GGGGTGAAACCCCGTCTTACTTAAATAACAAAAATCAGCGGGTGGAGTGGCGGCGC 1688
DB 345 --GGTGAACCCCGTCTTACTTAAATAACAAAAATAGCCGGGAGGTGGCGGCGC 288
QY 1689 CTGTAGTCCAGCTACTCAGGAGGTGAGGCAAGAGATGGGTGAACCCAGGCGGCGC 1748
DB 287 CTGTAGTCCAGCTACTCAGGAGGTGAGGCAAGAGATGGGTGAACCCAGGCGGCGC 228
QY 1749 AGCTACAGTACGCGAGATTCGCGCACTGCACTCAGCTGAGGAGATGAGTGGTGGTGGT 1808
DB 227 AGCTGCGAGTACGCGAGATTCGCGCACTGCACTCAGCTGAGGAGATGAGTGGTGGTGGT 168
QY 1809 TCTCAAAAAAAGAAAAAGAAAAAGAA 1838
DB 167 TCTCAAAAAAAGAAAAAGAAAAAGAA 138

RESULT 4
US-09-949-016-17189
; Sequence 17189, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17189
; LENGTH: 255679
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)---(255679)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17189

Query Match
Best Local Similarity 9.6%; Score 270.6; DB 4; Length 255679;
Matches 299; Conservative 1; Mismatches 28; Indels 2; Gaps 1;

```

```

Best Local Similarity 94.8%; Pred. No. 4.6e-68;
Matches 291; Conservative 0; Mismatches 14; Indels 2; Gaps 1;

QY 1532 GCCGGGCAACGGTGGCTCAGCCTGTAAATCCAGCACTTTGGGAGGCGGAGGCGGAGAT 1591
DB 230308 GCCGGGCGGGTGGCTCAGCCTGTAAATCCAGCACTTTGGGAGGCGGAGGCGGAGAT 230367
QY 1592 CACGAGGTGAGGAGATCGAGCATCTCTGGCTAAACACCGGGGTGAAACCCCGTCTTACTA 1651
DB 230368 CACGAGGTGAGGAGATCGAGCATCTCTGGCTAAACAC--GGTGAACCCCGTCTTACTA 230425
QY 1652 AAAATACAAAAAATCAGCGCGGTGAGGTGGCGGCGCCTGTAGTCCAGCTACTCAGGAG 1711
DB 230426 AAAATACAAAAAATTAACCGGCGGAGGTGGCGGCGCCTGTAGTCCAGCTACTCAGGAG 230485
QY 1712 GCTGAGGCAAGAGATGGCGTGAACCCAGGCGGCGGAGCCTTACAGTGAAGCGGAGATTGC 1771
DB 230486 GCTGAGGCAAGAGATGGCGTGAACCCAGGCGGCGGAGCCTTACAGTGAAGCGGAGATTGC 230545
QY 1772 GCCACTGCACTCCAGCCTGGAGCAGCAGTGAGACTCCGTCTCAAAAAAAGAAAAAGAA 1831
DB 230546 GCCACTGCACTCCAGCCTGGAGCAGCAGTGAGACTCCGTCTCAAAAAAAGAAAAAGAA 230605
QY 1832 AAAGAAA 1838
DB 230606 AAAAAA 230612

RESULT 5
US-09-949-016-16681/C
; Sequence 16681, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16681
; LENGTH: 51508
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16681

Query Match
Best Local Similarity 9.6%; Score 269.6; DB 4; Length 51508;
Matches 299; Conservative 0; Mismatches 29; Indels 2; Gaps 1;

QY 1509 AGACTGAAGAAAGCAGGAGTTGGCGGCACCGGTGGCTCAGCCTGTAAATCCAGCACT 1568
DB 12325 ATAATAATAATAAAAAAACAAGCGCGCGGATCAGAGGTGAGGAGATCAAGACCATCTGGCTAAACAC 12265
QY 1569 TTGGGAGGCGGAGGCGGCGGAGATCAGAGGTGAGGAGATCGAGCATCTCTGGCTAAACAC 1628
DB 12265 TTGGGAGGCGGAGGCGGCGGAGATCAGAGGTGAGGAGATCAAGACCATCTGGCTAAACAC 12206
QY 1629 GGGGTGAAACCCCGTCTTACTTAAATAACAAAAATCAGCGCGGTGAGGTGGCGGCGC 1688
DB 12205 --GGTGAACCCCGTCTTACTTAAATAACAAAAATAGCCGGGCAAGGTGGCGGCGC 12148
QY 1689 CTGTAGTCCAGCTACTCAGGAGGTGAGGCAAGAGATGGGTGAACCCAGGCGGCGC 1748
DB 12147 CTGTAGTCCAGCTACTCAGGAGGTGAGGCAAGAGATGGGTGAACCCAGGCGGCGC 12088

```

Qy	1749	AGCCTACGTAGCCGAGAT	TGGCCACATGCACT	CCAGCCTCGACGACGAGAGACTCGG	1808	
Db	12087	AGCCTGAGTGA	GCCGAGAT	TGGCCACATGCACT	CCAGCCTGGGCGACAGCAGAGACTCGG	12028
Qy	1809	TCTCAAAAAAAAAA	AGAAAGAAAAA	AGAAAA	1838	
Db	12027	TCTCAAAAAAAAAA	AAAAAAAAA	AAAAAACAA	11998	

RESULT 6
US-09-949-016-15535/c
; Sequence 15535, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15535
; LENGTH: 13132
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15535

```

; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16435
; LENGTH: 258775
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-16435

Query Match          9.5%; Score 267.4; DB 4; Length 258775;
Best Local Similarity 90.0%; Pred. No. 4.2e-67;
Matches 298; Conservative 0; Mismatches 31; Indels 2; Gaps 1;

Qy      1516 AGAAAAGCGCAGGAGTTCGCGGCGACGGTGCACGCTGTAATCCACAGCACTTTGGGAG 1575
Db      225045 AAATATCTTAGAAGAGCGCGGCGCGGTGTCATGCTGTAATCCACAGCACTTTGGAG 224986

Qy      1576 GCCGAGCGCGGCAGATCACGAGGTCAGGAGATCGAGACCATCTCTGGCTAAACACGGGGTCA 1635
Db      224985 GCCGAGTGGCGGATCACGAGGTCAGGAGATCGAGACCATCTCTGGCTAAACAC--AGTGA 224982

Qy      1636 AACCCCGTCTCTACTAAATAATACAAAATAACAGCCGGTGAGGTGGCGGGCGCCTGTAGT 1695
Db      224927 AACCCCGTCTCTACTAAATAATACAAAATAATAGCCGGCGAGGTGGCGGGCGCCTGTAGT 224868

Qy      1696 CCCAGCTACTCAGGAGGCTGAGGCAAGAGAAATGGCGTGAACCCACGAGGGGGCGGACCTTAC 1755
Db      224867 CCCAGCTACTCAGGAGGCTGAGGCAAGAGAAATGGCGTGAACCCACGAGGGGGCGGACCTTAC 224808

Qy      1756 AGTGAGCGGAGATTGCGCCACTGCACCTCCAGCCTGGAGCAGAGTGAGACTCGGTCTCAAA 1815
Db      224807 AGTGAGCGGAGATCGCGCCACTGCACCTCCAGCCTAGGCGCAGTGAGACTCGGTCTCAAAA 224748

Qy      1816 AAAAAAAAAAGAAAGAAAAAGGAGGAGGA 1846
Db      224747 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGA 224717

```

; APPLICANT: VENTER, J. CRAIG ET AL.
 ; TITLE OF INVENTION: POLYMORPHIS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016

Qy	1515	ARGAAGGACGAGNTTGGCGGCACGGTGCTCAGCCCTGTATTCACGACACTTTGGGA	1574
Db	18925	AAAAAAAAAAAAAGCCGGCGCGGTGGCTCACGCCCTGTAATCCCAGCACTTTGGGA	18766
Qy	1575	GGCCGAGCGGGCAGATCACGAGGTCAGGAGATCGAGACCATTCTGGCTAAACAGCGGGTG	1634
Db	18765	GGCCGAGCGGGTGGATCACGAGTCAAGGATCGAGACAATCTCTGGCTAACAC--GGTG	18708
Qy	1635	AATCCCGTCTCTACTATAAAAATCAAAAATCAGCCGGTGAGGTGGCGGGCGCCTGTAG	1694
Db	18707	AATCCCGTCTCTACTATAAAAATCAAAAATTAGCCGGCGAGGTGGCGGGCGCCTGTAG	18648
Qy	1695	TCCAGCTACTCAGGAGGCTGAGGCAAGAAATGGCGTGAAATCCAGCGGGCGGAGCCTA	1754
Db	18647	TCCTAGCTACTCCGAGGCTTGAGCAGGAAATGGCGTGAAATCCAGGAGGCGGAGCCTG	18588
Qy	1755	CAGTGACCGAGATTGGCCACTGCACTCCAGCTGGACGACAGTGAGACTTCCGCTCAA	1814
Db	18587	CAGTGACCGAGATTGGCCACTGCACTCCAGCTGGGCGACAGGAGACTCCGTATCAA	18528
Qy	1815	AAAAAAAAAAGAAAGAAAGAAAGCGAGGT	1848
Db	18527	AAAAAAAAAAGAAAGAAATTTGGACAGATCT	18494

```

RESULT 9
US-09-949-016-12234/c
; Sequence 12234, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIORITY FILING DATE: 60/241,755
; PRIORITY FILING DATE: 2000-10-20
; PRIORITY APPLICATION NUMBER: 60/237,768
; PRIORITY FILING DATE: 2000-10-03
; PRIORITY APPLICATION NUMBER: 60/231,498
; PRIORITY FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12234
; LENGTH: 91665
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12234

```

QY 1815 AAAAAAAAAAGAAAGAAAAAGGACGACT 1818
|||||
db 38026 AAAAAAAAAAGAAAGAAATTGGGACAGATCT 37993

RESULT 10

US-09-949-016-15829/c
; Sequence 15829, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15829
; LENGTH: 44166
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15829

Query Match 9.5%; Score 266.6; DB 4; Length 44166;
Best Local Similarity 90.5%; Pred. No. 2.1e-67;
Matches 296; Conservative 0; Mismatches 29; Indels 2; Gaps 1;

Qy	1507	TGAGACTCGAAGAAAAAGGCAGGAGTTGCGGGGCACGGTGGCTCAGCCCTGTAAATCCGACCA	1566
Db	25428	TCACACAAAGAAATCTCATATTTTGGCGGGCGGGTGGCTCAGCCCTGTAAATCCGACCA	25369
Qy	1567	CTTTGGAGGCGCGAGGGCGGAGATACGAGGTGAGGAGATCGAGACCATCTCGGTAAAC	1626
Db	25368	CTTTGGGAGGCGCAGGGCGGGTGGATCGAGGTGAGGAGATCGAGACCATCTCGGTAAAC	25309
Qy	1627	ACGGGGTGAACCCCGTCTCTACTTAAATAACAAAAATACGCGGGTGGAGTGGCGGGC	1686
Db	25308	AC--GGTGAACCCCGTCTCTACTTAAATAACAAAAATTAAGCCGGCGAGGTGGCGGGC	25251
Qy	1687	GCCTGTAGTCCACGTACTCAGGAGGCTGAGGCGAAGAGATGGCGGTGAACCCCGAGGGGC	1746
Db	25250	GCCTGTAGTCCACGTACTCAGGAGGCTGAGGCGAAGAGATGGTGTGAACCCCGAGGGGC	25191
Qy	1747	CGAGCTACAGTGAAGCGGAGATTCGGCGACTGCACTCAGCCCTGGACGACAGTGAGACTC	1806
Db	25190	AGAGCTTCGATGAAGCGGAGATTCGGCGACTGCACTCAGCCCTGGCGACGACGAGACTC	25131
Qy	1807	CGTCTCAAAAAAAGAAAGAAAGAA 1833	
Db	25130	CGTCTCAAAAAAAGAAAGAAAGAA 25104	

RESULT 11

US-09-949-016-14894/c
; Sequence 14894, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

```
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14894
; LENGTH: 54180
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(54180)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14894

Query Match
Best Local Similarity 9.5%; Score 266.6; DB 4; Length 54180;
Matches 296; Conservative 0; Mismatches 29; Indels 2; Gaps 1;

QY 1507 TGAGACTGAAGAAAGGAGGAGTTGCCGGGACCGTGGCTCAGCGCTGTATATCCAGCA 1566
Db 35582 TCACACAAAGAATCTCATATTTTGGCGGGCGGGTGGCTCAGCGCTGTATATCCAGCA 35523
QY 1567 CTTTGGGAGCGCGAGCGGGGAGATCAGAGTCAAGAGATCGAGACCATCTGGCTAAC 1626
Db 35522 CTTTGGGAGCGCGAGCGGGTGGATCAGAGTCAAGAGATCGAGACCATCTGGCTAAC 35463
QY 1627 ACGGGGTGAACCCCGTCTCTACTTAAATAACAAAAATACAAAAATACGCGGGTGAGGTGGCGGC 1686
Db 35462 AC--GGTGAACCCCGTCTCTACTTAAATAACAAAAATAGCCGGCGAGTGGCGGC 35405
QY 1687 GCCTGTAGTCCCGTACTCAGGAGCTGAGCAGAGAGAGTGGGTGAACCCAGGGGC 1746
Db 35404 GCCTGTAGTCCCGTACTCGGAGGCTGAGGAGAGAGTGGGTGAACCCAGGGGC 35345
QY 1747 CGAGCTACAGTGAAGCCAGAGATTGGCCACTGCATCTCCAGCTGAGCAGCAGTGAAGTCTC 1806
Db 35344 AGACCTGCATGAGCCAGAGTTGGCCACTGCATCTCCAGCTGAGCAGCAGTGAAGTCTC 35285
QY 1807 CGTCTCAAAAAAAGAAAGAAAGAA 1833
Db 35284 CGTCTCAAAAAAAGAAAGAAAGAA 35258

RESULT 12
US-09-949-016-11804
; Sequence 11804, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11804
; LENGTH: 67479
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(67479)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11804
```

```
Query Match
Best Local Similarity 9.4%; Score 266.4; DB 4; Length 67479;
Matches 297; Conservative 0; Mismatches 31; Indels 2; Gaps 1;

QY 1511 ACTGAAGAAAGGAGGAGTTGCCGGGACCGTGGCTCAGCGCTGTATATCCAGCACTTT 1570
Db 33321 AATTTATAAGAAAAACAGGTGGCGGGCGGTGGCTCAGCGCTGTATATCCAGCACTTT 33380
QY 1571 GGGAGGCGGAGCGGGCAGATCAGAGTCAAGAGATCGAGACCATCTCTGGCTAACAGG 1630
Db 33381 GGGAGGCGGAGCGGGCGGATCAGAGTCAAGAGATCGAGACCATCTCTGGCTAACAC-- 33438
QY 1631 GGTGAACCCCGTCTCTACTTAAATAACAAAAATACAAAAATACGCGGGTGAGGTGGCGGCCT 1690
Db 33439 GGTGAACCCCGTCTCTACTTAAATAACAAAAATACAAAAATACGCGGGTGAGGTGGCGGCCT 33498
QY 1691 GTAGTCCCAGTACTCAGGAGGCTGAGGCAAGAGAATGGCGTGAACCCAGGGGCCGAG 1750
Db 33499 GTAGTCCCAGTACTCAGGAGGCTGAGGCAAGAGAATGGCGTGAACCCAGGGGCCGAG 33558
QY 1751 CCTACAGTGAAGCCAGATTGGCCACTGCATCTCAGCGCTGAGCAGCAGTGAAGTCTCCCTC 1810
Db 33559 GCTCAGTGAAGCCAGATTGGCCACTGCATCTCAGCGTGGGCGCAGAGAGTCTCCCTC 33618
QY 1811 TCAAAAAAAGAAAGAAAGAAAGAAAG 1840
Db 33619 TCAAAAAAAGAAAGAAAGAAAGAAAG 33648

RESULT 13
US-09-949-016-15358
; Sequence 15358, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15358
; LENGTH: 71119
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-15358

Query Match
Best Local Similarity 9.4%; Score 266.4; DB 4; Length 71119;
Matches 297; Conservative 0; Mismatches 31; Indels 2; Gaps 1;

QY 1511 ACTGAAGAAAGGAGGAGTTGCCGGGACCGTGGCTCAGCGCTGTATATCCAGCACTTT 1570
Db 11100 AATTTATAAGAAAAACAGGTGGCGGGCGGTGGCTCAGCGCTGTATATCCAGCACTTT 11159
QY 1571 GGGAGGCGGAGCGGGCAGATCAGAGTCAAGAGATCGAGACCATCTCTGGCTAACAGG 1630
Db 11160 GGGAGGCGGAGCGGGCGGATCAGAGTCAAGAGATCGAGACCATCTCTGGCTAACAC-- 11217
QY 1631 GGTGAACCCCGTCTCTACTTAAATAACAAAAATACAAAAATACGCGGGTGAGGTGGCGGCCT 1690
Db 11218 GGTGAACCCCGTCTCTACTTAAATAACAAAAATACAAAAATACGCGGGTGAGGTGGCGGCCT 11277
QY 1691 GTAGTCCCAGTACTCAGGAGGCTGAGGCAAGAGAATGGCGTGAACCCAGGGGCCGAG 1750
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2005, 01:06:08 ; Search time 1103.5 Seconds
(without alignments)
15978.923 Million cell updates/sec

Title: US-08-731-499-6

Perfect score: 2821

Sequence: 1 ATCTTAGACGACAGCCTG.....AGGGTTTCAGACGAGCCTG 2821

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 6067389 seqs, 3125258755 residues

Total number of hits satisfying chosen parameters: 12134778

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA:*

1: /cgm2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgm2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgm2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgm2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgm2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgm2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgm2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgm2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgm2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgm2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgm2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgm2_6/ptodata/1/pubpna/US09D_PUBCOMB.seq.*
13: /cgm2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgm2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgm2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgm2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgm2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgm2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgm2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgm2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgm2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgm2_6/ptodata/1/pubpna/US10J_PUBCOMB.seq.*
23: /cgm2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgm2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgm2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgm2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2821	100.0	2821	28	US-08-731-499-6
2	1103.4	39.1	11006	22	Sequence 6, Appli
3	1103.4	39.1	11006	22	Sequence 11, Appl
4	559.6	19.8	600	13	Sequence 11, Appl
5	559.6	19.8	600	17	Sequence 134687
6	337	11.9	1506	20	Sequence 134687
7	337	11.9	1506	20	Sequence 23318, A
					Sequence 29199, A

ALIGNMENTS

RESULT 1

US-08-731-499-6
; Sequence 6, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION: Joe W.
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOMBEL, David
; APPLICANT: ROMMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,499
; FILING DATE: 16-OCT-1996
; CLASSIFICATION: 435

Sequence 39, Appl
Sequence 328, App
Sequence 518, App
Sequence 328, App
Sequence 1240, App
Sequence 840, App
Sequence 544, App
Sequence 5, Appli
Sequence 16, Appl
Sequence 16, Appl
Sequence 12, Appl
Sequence 36, Appl
Sequence 665, App
Sequence 5775, App
Sequence 5798, App
Sequence 250, App
Sequence 261000,
Sequence 261000,
Sequence 1, Appli
Sequence 17630, A
Sequence 195045,
Sequence 195046,
Sequence 195045,
Sequence 195046,
Sequence 5797, App
Sequence 5799, App
Sequence 10, Appl
Sequence 750, App
Sequence 82, Appl
Sequence 9556, App
Sequence 884, App
Sequence 172, App
Sequence 4, Appli
Sequence 258046,
Sequence 258046,
Sequence 10219, A
Sequence 2197, App
Sequence 2197, App

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/680,395
; FILING DATE: 15-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 23070-068910
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2821 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: 1..2821 /note= "cDNA clone GCAP encodes a
; OTHER INFORMATION: guanino cyclase activating protein
; OTHER INFORMATION:
;
US-08-731-499-6

Query Match 100.0%; Score 2821; DB 8; Length 2821;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2821; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCCTAAGACGACACAGCCTGGGAAGCCAGACACTGGGGAAGTGGTCTCAGGGGATGTGGGT 60
Db 1 ATCCTAAGACGACACAGCCTGGGAAGCCAGACACTGGGGAAGTGGTCTCAGGGGATGTGGGT 60

Qy 61 CACTGGGGTGAAGTGGAGCTTTTCAGGGTCTCCCGTCAATGCAGCTGAGTTTCTTTGGC 120
Db 61 CACTGGGGTGAAGTGGAGCTTTTCAGGGTCTCCCGTCAATGCAGCTGAGTTTCTTTGGC 120

Qy 121 AGGGAATTTACAGCTGAAGAAAGCTCCCGCGGAGACTACAACTCAGCAAGCCAGC 180
Db 121 AGGGAATTTACAGCTGAAGAAAGCTCCCGCGGAGACTACAACTCAGCAAGCCAGC 180

Qy 181 TGCTCACACCCGAGAGTCTGGACAGGATCTTCTCTCTGGTGGATGAGAGATG 240
Db 181 TGCTCACACCCGAGAGTCTGGACAGGATCTTCTCTCTGGTGGATGAGAGATG 240

Qy 241 GTAAGAGGGGAGAGATGGGAGAGTGTCTTCCACTCTGCATCATCGCCACTTTCTGGCC 300
Db 241 GTAAGAGGGGAGAGATGGGAGAGTGTCTTCCACTCTGCATCATCGCCACTTTCTGGCC 300

Qy 301 GCAGTCTCTGGGCAAGGCCCTCCACCTTCCAACTGGGGTCTCATCTGTGAGAGGC 360
Db 301 GCAGTCTCTGGGCAAGGCCCTCCACCTTCCAACTGGGGTCTCATCTGTGAGAGGC 360

Qy 361 TGTGAGAAGATGTGATGAACCTAAAGAGGACTCATGAGCAGCGTGTGTGAGAGTGAC 420
Db 361 TGTGAGAAGATGTGATGAACCTAAAGAGGACTCATGAGCAGCGTGTGTGAGAGTGAC 420

Qy 421 TAAAGTCTTACAGAGTTGCTGATGGAGGCCAGGACGCGAGATAGAAAGATAGGAAC 480
Db 421 TAAAGTCTTACAGAGTTGCTGATGGAGGCCAGGACGCGAGATAGAAAGATAGGAAC 480

Qy 481 TTTGGAGTCAGCGAGGAGTGCATATATTAGCTTCTCGTCTAGTCTCAATTTCTCATC 540
Db 481 TTTGGAGTCAGCGAGGAGTGCATATATTAGCTTCTCGTCTAGTCTCAATTTCTCATC 540

Qy 541 TGGAAATGGGGATAAATAGTGGTTGAGAGGAATGAATAGGATAATGTGTTTAAAGAGC 600
Db 541 TGGAAATGGGGATAAATAGTGGTTGAGAGGAATGAATAGGATAATGTGTTTAAAGAGC 600

Qy 601 AGGCATAGGGTAGACCTTCATTCAGGCTGCTTGGGCTTTCTCCCTGTAGGCCCAAGGCC 660
Db 601 AGGCATAGGGTAGACCTTCATTCAGGCTGCTTGGGCTTTCTCCCTGTAGGCCCAAGGCC 660

```

```

661 AGCCTCAGGGCTATGTGGGAGAGAGCTGGCTTGGAAATACACACTTGAGCCCTCCAGCTC 720
Db 661 AGCCTCAGGGCTATGTGGGAGAGAGCTGGCTTGGAAATACACACTTGAGCCCTCCAGCTC 720

Qy 721 TCTCAGCTCCACCAGCATTTCCGTGTPACCATCGCAAAAGTAAACTTCAATTCATCA 780
Db 721 TCTCAGCTCCACCAGCATTTCCGTGTPACCATCGCAAAAGTAAACTTCAATTCATCA 780

Qy 781 GCAAAGAAAGCCCTTAAAGTGCAGGAGACTCTGGAGATTGAGACACCTGCAAGCC 840
Db 781 GCAAAGAAAGCCCTTAAAGTGCAGGAGACTCTGGAGATTGAGACACCTGCAAGCC 840

Qy 841 GCAAGCTTGAGTCTGAGACTGCAAGGATAGTTGGCATAAAGACGTGTAGCGCATCCTGGG 900
Db 841 GCAAGCTTGAGTCTGAGACTGCAAGGATAGTTGGCATAAAGACGTGTAGCGCATCCTGGG 900

Qy 901 AGCAGGCTCTCTCTCTGCCCCCAGACCCAGGCTCTCCCTTCTTCTATACATGACCACTC 960
Db 901 AGCAGGCTCTCTCTCTGCCCCCAGACCCAGGCTCTCCCTTCTTCTATACATGACCACTC 960

Qy 961 TCCTCCCCCTTGCTCAGGCCAGCTGTCTCTGAAAGAGTTTCTGAAAGTGGCCGCTCGGA 1020
Db 961 TCCTCCCCCTTGCTCAGGCCAGCTGTCTCTGAAAGAGTTTCTGAAAGTGGCCGCTCGGA 1020

Qy 1021 CAAGTGGGTGATGAAGATGCTGCAGATGACATGAATCCAGCAGCTGCTCGCTCAGCA 1080
Db 1021 CAAGTGGGTGATGAAGATGCTGCAGATGACATGAATCCAGCAGCTGCTCGCTCAGCA 1080

Qy 1081 GAGACGGAAGAGTCCATGTTCTGAGGAGTCTGGGGCCCCCTCCACGACTCCAGGCTCACC 1140
Db 1081 GAGACGGAAGAGTCCATGTTCTGAGGAGTCTGGGGCCCCCTCCACGACTCCAGGCTCACC 1140

Qy 1141 CAGGTTTCCAGGTAAGTGGAGGCTCCCTGGCTCAGCTGCTCATGCCCACTCTTCCCC 1200
Db 1141 CAGGTTTCCAGGTAAGTGGAGGCTCCCTGGCTCAGCTGCTCATGCCCACTCTTCCCC 1200

Qy 1201 TGGTGTGACTTCTCTGGCACCCTCTGTCAGGGCTGAGTGGGGATGGGGAAGGGCTGTG 1260
Db 1201 TGGTGTGACTTCTCTGGCACCCTCTGTCAGGGCTGAGTGGGGATGGGGAAGGGCTGTG 1260

Qy 1261 GGTGTTGAAGTGGCCAAAGGAGCTAGTCCATTTTGGAGGAGTCCCTGGGATGGTGAAGG 1320
Db 1261 GGTGTTGAAGTGGCCAAAGGAGCTAGTCCATTTTGGAGGAGTCCCTGGGATGGTGAAGG 1320

Qy 1321 AATTTCAGTTACTTTTCTGTTTTCAGCGCTCTCTGGGAGACTGTGCTTTGGCTGGGTGTT 1380
Db 1321 AATTTCAGTTACTTTTCTGTTTTCAGCGCTCTCTGGGAGACTGTGCTTTGGCTGGGTGTT 1380

Qy 1381 GTGGGGCTCCCACTAGTTTCTGGGTGTTTCTCAAGTGGAGCAAGAGCCAACTGAGGGGTGA 1440
Db 1381 GTGGGGCTCCCACTAGTTTCTGGGTGTTTCTCAAGTGGAGCAAGAGCCAACTGAGGGGTGA 1440

Qy 1441 GGGTCCCAAGCAAAATCAGAAATGAGAACACAAAGACTGGTAGGAGGCAAGGGGTGGGA 1500
Db 1441 GGGTCCCAAGCAAAATCAGAAATGAGAACACAAAGACTGGTAGGAGGCAAGGGGTGGGA 1500

Qy 1501 GGGTGTGAGACTGAAAGAAAGGAGGTTGCGGGCAGCGTGGCTCACCGCTGTAATC 1560
Db 1501 GGGTGTGAGACTGAAAGAAAGGAGGTTGCGGGCAGCGTGGCTCACCGCTGTAATC 1560

Qy 1561 CCAGCACTTTGGGAGGCCGAGCGGGGAGATCAGAGGTGAGGAGATCGAGACCATCTCTG 1620
Db 1561 CCAGCACTTTGGGAGGCCGAGCGGGGAGATCAGAGGTGAGGAGATCGAGACCATCTCTG 1620

Qy 1621 GCTAACACGGGGTGAACCCCGTCTCTAATAAATAAATAATACAGCCCGGGTGAAGTG 1680
Db 1621 GCTAACACGGGGTGAACCCCGTCTCTAATAAATAAATAATACAGCCCGGGTGAAGTG 1680

Qy 1681 GCGGCGCCTCTAGTCCAGCTACTCAGGAGGCTGAGCAAGAGAAAGCGCTGAAACCCCA 1740
Db 1681 GCGGCGCCTCTAGTCCAGCTACTCAGGAGGCTGAGCAAGAGAAAGCGCTGAAACCCCA 1740

Qy 1741 GGGGGCCGAGCTTACAGTGAAGCGAGATGTCGCCCACTGCACTCCAGCCTGGACGAGT 1800

```

Db 1741 GGGGGCCGAGCCTACAGTGAGCGAGATTGCGCACTGCACTCCAGCCTGGACACAGTG 1800
Qy 1801 AGACTCCGCTCTCAAAAAAAGAAAAAGAAAAAGGAGGAGGAGTTTGGGGGGCAG 1860
Db 1801 AGACTCCGCTCTCAAAAAAAGAAAAAGAAAAAGGAGGAGGAGTTTGGGGGGCAG 1860
Qy 1861 GGGGAGCAGTAATTTCTATACTTCGCGGATGCTGAGGGCGGTTTCATGGGAGGACCCCTG 1920
Db 1861 GGGGAGCAGTAATTTCTATACTTCGCGGATGCTGAGGGCGGTTTCATGGGAGGACCCCTG 1920
Qy 1921 GCCTCTCTCCCAAGGATCTCCACAGTGTGTCAACAGGAAAAATGGCAGCAATA 1980
Db 1921 GCCTCTCTCCCAAGGATCTCCACAGTGTGTCAACAGGAAAAATGGCAGCAATA 1980
Qy 1981 CGCTGAGGCTGTGCTTTCTGCTTTGAAAGGGTCAGCTGTACTTTAAAGGACTGTGTT 2040
Db 1981 CGCTGAGGCTGTGCTTTCTGCTTTGAAAGGGTCAGCTGTACTTTAAAGGACTGTGTT 2040
Qy 2041 CAGCTCTGCTGGTGTCTGCTGGGACCCCTGCTGCCAACCCACCACTCTCCCCCAACAA 2100
Db 2041 CAGCTCTGCTGGTGTCTGCTGGGACCCCTGCTGCCAACCCACCACTCTCCCCCAACAA 2100
Qy 2101 TCCTCTCTTCCATCCATATCCCAAGTATGACCTTCCACACTCCAGCCATAGCTG 2160
Db 2101 TCCTCTCTTCCATCCATATCCCAAGTATGACCTTCCACACTCCAGCCATAGCTG 2160
Qy 2161 AATGTTTCTTTAAAGGATGGAGAAAACTTCTGCTGTCTCTGGCAAGAAATTTGGGGAC 2220
Db 2161 AATGTTTCTTTAAAGGATGGAGAAAACTTCTGCTGTCTCTGGCAAGAAATTTGGGGAC 2220
Qy 2221 TGTGACTGGGATTTGGGCTGGGCTGCTTAACTGCTGTGTGAACCCACAGACGCCA 2280
Db 2221 TGTGACTGGGATTTGGGCTGGGCTGCTTAACTGCTGTGTGAACCCACAGACGCCA 2280
Qy 2281 CTTCTCTCCCTAACTTGTATGCTTGGCAGACAGTGAGCGTGGGACTAGGCGA 2340
Db 2281 CTTCTCTCCCTAACTTGTATGCTTGGCAGACAGTGAGCGTGGGACTAGGCGA 2340
Qy 2341 ACAGTTTGGATATTTGCTTTAGATGTGAATTTATTTTGTATATAAATCTTTAT 2400
Db 2341 ACAGTTTGGATATTTGCTTTAGATGTGAATTTATTTTGTATATAAATCTTTAT 2400
Qy 2401 GTGTAAACCCCAATATAGAACTAGATTAAGAGGAGTCTCTCTGTTGAAAGGGGAGCTG 2460
Db 2401 GTGTAAACCCCAATATAGAACTAGATTAAGAGGAGTCTCTCTGTTGAAAGGGGAGCTG 2460
Qy 2461 AGTACCTCTGAACTGGAGGACCTCTGAAAGAGCAAACTGAAACCCAGTGCCTTGG 2520
Db 2461 AGTACCTCTGAACTGGAGGACCTCTGAAAGAGCAAACTGAAACCCAGTGCCTTGG 2520
Qy 2521 TCACCTGTTACTCTATAAGACAGTTTAAAGTGAGACCTGGAAGAAACATTTTACCTT 2580
Db 2521 TCACCTGTTACTCTATAAGACAGTTTAAAGTGAGACCTGGAAGAAACATTTTACCTT 2580
Qy 2581 GAATAGATAGGTTTATGTTGATATAGAAATAAACTAACCTATTAAACCCCTGAGAC 2640
Db 2581 GAATAGATAGGTTTATGTTGATATAGAAATAAACTAACCTATTAAACCCCTGAGAC 2640
Qy 2641 TTTCAGGTTGTTTATTTATATATAGTATATATAAAATTTCTTTAGACATCAATTTTA 2700
Db 2641 TTTCAGGTTGTTTATTTATATATAGTATATATAAAATTTCTTTAGACATCAATTTTA 2700
Qy 2701 GGTAAAAAATATTTAGTAAAAAATTTTGGCCAGGTGAGCAGCTCACACCTGCAATCC 2760
Db 2701 GGTAAAAAATATTTAGTAAAAAATTTTGGCCAGGTGAGCAGCTCACACCTGCAATCC 2760
Qy 2761 CAGGACTTTGGAGGCCGAGGGGTGATCACTGAGTCAAGGGTTCAAGACCAAGCCT 2820
Db 2761 CAGGACTTTGGAGGCCGAGGGGTGATCACTGAGTCAAGGGTTCAAGACCAAGCCT 2820
Qy 2821 G 2821

Db 2821 G 2821
RESULT 2
US-10-737-082-11
; Sequence 11, Application US/10737082
; Publication No. US20050130170A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Beard, Chris
; APPLICANT: Burgess, Chris
; APPLICANT: Gannon, Allison
; APPLICANT: Harvey, Jeanne
; APPLICANT: Lechner, John F.
; APPLICANT: Li, Zheng
; TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
; FILE REFERENCE: 1657/2032
; CURRENT APPLICATION NUMBER: US/10/737,082
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: US 10/737,082
; PRIOR FILING DATE: 2003-12-16
; NUMBER OF SEQ ID NOS: 300
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 11006
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-737-082-11
Query Match 39.1%; Score 1103.4; DB 22; Length 11006;
Best Local Similarity 99.9%; Pred. No. 1.4e-299;
Matches 1104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATCTTAAGACGACAGCCTGGGAAGCCAGCACTGGGGAAGTGTGTCTGAGGAGTGTGGGT 60
Db 9902 ATCTTAAGACGACAGCCTGGGAAGCCAGCACTGGGGAAGTGTGTCTGAGGAGTGTGGGT 9961
Qy 61 CACTGGGTGAAGGTGGAGCTTTCAGGGTCTCCCGTCAATGACAGCTGAGTTTCTTTGGC 120
Db 9962 CACTGGGTGAAGGTGGAGCTTTCAGGGTCTCCCGTCAATGACAGCTGAGTTTCTTTGGC 10021
Qy 121 AGGGAATTTACAGCTGAAGAAAGCTGCGGCGAGAGCTACAACTGACAGAGCCAGC 180
Db 10022 AGGGAATTTACAGCTGAAGAAAGCTGCGGCGAGAGCTACAACTGACAGAGCCAGC 10081
Qy 181 TGCTCACACCCGAGGAGGTCTGGACAGGATCTTCTCTGTTGGATGAGAAATGGAGATG 240
Db 10082 TGCTCACACCCGAGGAGGTCTGGACAGGATCTTCTCTGTTGGATGAGAAATGGAGATG 10141
Qy 241 GTAAGAGGGCAGAGATGGGAGAGTGTCTCCACTCTGCATCATCGCCACTTTCTGGCC 300
Db 10142 GTAAGAGGGCAGAGATGGGAGAGTGTCTCCACTCTGCATCATCGCCACTTTCTGGCC 10201
Qy 301 GCAGTCTCTGGGCAAGGCCCTCCACCTTCCAACTGGGGTCTCATCTGTGAGAGGC 360
Db 10202 GCAGTCTCTGGGCAAGGCCCTCCACCTTCCAACTGGGGTCTCATCTGTGAGAGGC 10261
Qy 361 TGTGAGAGAGATGTCTATGAATTAACAAAGGGAGCTCATGAGCAGCTGTTGTAGAGTGAC 420
Db 10262 TGTGAGAGAGATGTCTATGAATTAACAAAGGGAGCTCATGAGCAGCTGTTGTAGAGTGAC 10321
Qy 421 TAAAGTCTTACAGGAGTTGCTGATGGAGCCAGGCCAGCAATAGAAATAGAAAC 480
Db 10322 TAAAGTCTTACAGGAGTTGCTGATGGAGCCAGGCCAGCAATAGAAATAGAAAC 10381
Qy 481 TTTGGAGTCAGCAGGAGTGATATTTGAGCTTCTCGTCTTAGTCTCAATTTCTTCATC 540
Db 10382 TTTGGAGTCAGCAGGAGTGATATTTGAGCTTCTCGTCTTAGTCTCAATTTCTTCATC 10441
Qy 541 TCGAAAAATGGGATTAATAATAGTGGTTGAGAGGAATGAATAGGATTAATGTGTTAAGAGC 600
Db 10442 TCGAAAAATGGGATTAATAATAGTGGTTGAGAGGAATGAATAGGATTAATGTGTTAAGAGC 10501

Qy	601	AGGCATAGGGT	GAGACCT	CCATT	CAGCGTG	CTTTGGG	CTTTCC	TCCCT	GTAGCC	CAAGCC	660
Db	10502	AGGCATAGGGT	GAGACCT	CCATT	CAGCGTG	CTTTGGG	CTTTCC	TCCCT	GTAGCC	CAAGCC	10561
Qy	661	AGCCTCAGGG	CTATGTGGG	GAGAGAC	TGGCT	TGGNA	TACAC	ACTT	GAGCC	CTCAG	720
Db	10562	AGCCTCAGGG	CTATGTGGG	GAGAGAC	TGGCT	TGGNA	TACAC	ACTT	GAGCC	CTCAG	10621
Qy	721	TCTCAG	CTCCACC	CAGCAT	TTTCCG	TGCTAC	CATGCG	CAAAAG	TAAAA	CTTCA	780
Db	10622	TCTCAG	CTCCACC	CAGCAT	TTTCCG	TGCTAC	CATGCG	CAAAAG	TAAAA	CTTCA	10681
Qy	781	GCAAGAA	AGCCCTTAA	AGGTGG	CAGGAG	ACTCT	CGAGAT	TTCAG	AACACT	GAACGC	840
Db	10682	GCAAGAA	AGCCCTTAA	AGGTGG	CAGGAG	ACTCT	CGAGAT	TTCAG	AACACT	GAACGC	10741
Qy	841	GCAAGCT	TGAGGTCT	GAGACT	GTGAG	GATAG	TGGCA	TAAGA	CGTGT	PAGCG	900
Db	10742	GCAAGCT	TGAGGTCT	GAGACT	GTGAG	GATAG	TGGCA	TAAGA	CGTGT	PAGCG	10801
Qy	901	AGCGAGG	CTCTCCT	CCTGCC	CCAC	AGCCAG	GTCTCC	CCCTT	CTTCT	TACAT	960
Db	10802	AGCGAGG	CTCTCCT	CCTGCC	CCAC	AGCCAG	GTCTCC	CCCTT	CTTCT	TACAT	10861
Qy	961	TCCT	CCCCCT	TGCTCAGG	CCAG	CTGTCT	TGAAC	GAGTTT	GTGTAAG	GTGCC	1020
Db	10862	TCCT	CCCCCT	TGCTCAGG	CCAG	CTGTCT	TGAAC	GAGTTT	GTGTAAG	GTGCC	10921
Qy	1021	CAAGTGGG	TGATGA	AGATG	CTGCAG	ATGG	ACAT	TGAAT	CCAC	CAGCTGG	1080
Db	10922	CAAGTGGG	TGATGA	AGATG	CTGCAG	ATGG	ACAT	TGAAT	CCAC	CAGCTGG	10981
Qy	1081	GAGACG	AAAGTG	CCATG	TTCT	TGA	1105				
Db	10982	GAGACG	AAAGTG	CCATG	TTCT	TGA	11006				

RESULT 3

```

US-10-765-790-11
; Sequence 11, Application US/10765790
; Publication No. US20050130172A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Beard, Chris
; APPLICANT: Burgess, Chris
; APPLICANT: Gannon, Allison
; APPLICANT: Harvey, Jeanne
; APPLICANT: Lechner, John F.
; APPLICANT: Li, Zheng
; TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
; FILE REFERENCE: 1657/2035
; CURRENT APPLICATION NUMBER: US/10/765,790
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US 10/737,082
; PRIOR FILING DATE: 2003-12-16
; NUMBER OF SEQ ID NOS: 300
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 11006
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-765-790-11

```

	Query Match	39.1%	Score 1103.4;	DB 22;	Length 11006;
	Best Local Similarity	99.9%	Pred. No. 1.4e-239;		
	Matches 1104;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1	ATCCTAGAGCGCAGAGCTGGGAAGCAGACATGGGGAAGTGGTGTGACGGGATGTGGGT	60		
Db	9902	ATCCTAGAGCGCAGAGCTGGGAAGCAGACATGGGGAAGTGGTGTGACGGGATGTGGGT	9961		
Qy	61	CACCTGGGGTGAAGGTGGAGCTTTTCAGGGTCTCCGTCAAATGCAGCTCAGTTTTTCTTTGGC	120		

9962	Db	CACTGGGGTGAAGGTGGAGGCTTTTCAGGGTCTCCGGTCAATGCAGCTGAGTGTTCCTTTGGC	10021
121	Qy	AGGGAATTTTACCAGCTGAAGAAGCCCTGCGGCGAGAGCTACAAACTGAGCAAGGCCAGC	180
10022	Db	AGGGAATTTTACCAGCTGAAGAAGCCCTGCGGCGAGAGCTACAAACTGAGCAAGGCCAGC	10081
181	Qy	TGCTCACACCCGAGGAGTGTGGACAGGATCTTCTCTCTGTGTGATGAGAAATGAGATG	240
10082	Db	TGCTCACACCCGAGGAGTGTGGACAGGATCTTCTCTCTGTGTGATGAGAAATGAGATG	10141
241	Qy	GTAAGAGGGGCAGAGATGGGAGAGTGTCCACTCTGCAATCAGCCACATCTTCGGCC	300
10142	Db	GTAAGAGGGGCAGAGATGGGAGAGTGTCTCACTCTGCATCATGCCCACTTTCGGCC	10201
301	Qy	GCACGTCTTTGGGCAAGGCCCTCCACCTTCCAACCTGGGGTCCCTCATCTGTGAGAAGGC	360
10202	Db	GCACGTCTTTGGGCAAGGCCCTCCACCTTCCAACCTGGGGTCCCTCATCTGTGAGAAGGC	10261
361	Qy	TGTGGAGAGATGTCATGAACATAAAGGGACTCATGACGACGTGTTGTAGGAGTGAC	420
10262	Db	TGTGGAGAGATGTCATGAACATAAAGGGACTCATGACGACGTGTTGTAGGAGTGAC	10321
421	Qy	TAAAGTCTCTACAGGAGTTTGCTGATGGAGCCAGGCACGCAAGATAGAAAGAAATAGGAAC	480
10322	Db	TAAAGTCTCTACAGGAGTTTGCTGATGGAGCCAGGCACGCAAGATAGAAAGAAATAGGAAC	10381
481	Qy	TTTGGAGTCAGGCAGGAGTGATATATGAGCTTCTCGTCTAGTCTCAATTTCTCTCATC	540
10382	Db	TTTGGAGTCAGGCAGGAGTGATATATGAGCTTCTCGTCTAGTCTCAATTTCTCTCATC	10441
541	Qy	TGCAAATGGGGATAATAATAGTGGTTGAGAGGAATGAATAGATAATGCTGTTTAAAGAC	600
10442	Db	TGCAAATGGGGATAATAATAGTGGTTGAGAGGAATGAATAGATAATGCTGTTTAAAGAC	10501
601	Qy	AGGCATAGGCTAGACCTCCAATTCAGGCTGCTTTGGGCTTTCTCTCTGTAGCCCAAGCCC	660
10502	Db	AGGCATAGGCTAGACCTCCAATTCAGGCTGCTTTGGGCTTTCTCTCTGTAGCCCAAGCCC	10561
661	Qy	AGCCTCAGGGCTATGTGGGAGAGAGCTGGCTTGGATACACACTTGAGCCCTCCAGCTC	720
10562	Db	AGCCTCAGGGCTATGTGGGAGAGAGCTGGCTTGGATACACACTTGAGCCCTCCAGCTC	10621
721	Qy	TCTCAGCTCCACCCAGCATTTCCGTGTACATGGGCAAAAGTAAACTTCAATTCATCA	780
10622	Db	TCTCAGCTCCACCCAGCATTTCCGTGTACATGGGCAAAAGTAAACTTCAATTCATCA	10681
781	Qy	GCAAGAAAGCCCTTTAAAGTGGCAGGAGATCTCTGGAGATTCAGACACTGCAAGCC	840
10682	Db	GCAAGAAAGCCCTTTAAAGTGGCAGGAGATCTCTGGAGATTCAGACACTGCAAGCC	10741
841	Qy	GCAAGCTTGAGTCTCAGACTCGAGATAGTTGGCATAGAGCTGTAGGCGCATCCTGG	900
10742	Db	GCAAGCTTGAGTCTCAGACTCGAGATAGTTGGCATAGAGCTGTAGGCGCATCCTGG	10801
901	Qy	AGCGAGGTCTCTCCTCTGCCCCAGACCCAGGTCTCCCTTTCTTCTACATGACCACTC	960
10802	Db	AGCGAGGTCTCTCCTCTGCCCCAGACCCAGGTCTCCCTTTCTTCTACATGACCACTC	10861
961	Qy	TCCTCCCTCTGCTCAGGCAGCTGTCTCTGAACGAGTTTGTGAAGGTGCCCGTCGGGA	1020
10862	Db	TCCTCCCTCTGCTCAGGCAGCTGTCTCTGAACGAGTTTGTGAAGGTGCCCGTCGGGA	10921
1021	Qy	CAAGTGGGTGATGAAGATGCTGCAGATGACATGAATCCAGCAGCTGGCTCGCTCAGCA	1080
10922	Db	CAAGTGGGTGATGAAGATGCTGCAGATGACATGAATCCAGCAGCTGGCTCGCTCAGCA	10981
1081	Qy	GAGACGAAAGTGCCATGTTCTGA	1105
10982	Db	GAGACGAAAGTGCCATGTTCTGA	11006

; Sequence 134687, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134687
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-134687

Query Match 19.8%; Score 559.6; DB 13; Length 600;
Best Local Similarity 97.5%; Pred. No. 8e-147;
Matches 589; Conservative 1; Mismatches 10; Indels 4; Gaps 2;

QY 1068 GGCTGCTCAGCAGACGCGAAAGTGCATGTTCTGAGGAGTCTGGGGCCCTCCACGA 1127
DB 1 GGCTGCTCAGCAGACGCGAAAGTGCATGTTCTGAGGAGTCTGGGGCCCTCCACGA 60

QY 1128 CTCACGGCTCACCAGGTTTCCAGGTTAGTGGGTCCTGGCTCAGGCTGCTCATG 1187
DB 61 CTCACGGCTCACCAGGTTTCCAGGTTAGTGGGTCCTGGCTCAGGCTGCTCATG 118

QY 1188 CCCACTCTTCCCTGCTGTTGACTTCTCTGGCACCCCTCTGTCAGGCTGAGTGGGATG 1247
DB 119 CCCACTCTTCCCTGCTGTTGACTTCTCTGGCACCCCTCTGTCAGGCTGAGTGGGATG 178

QY 1248 GGAAGGCTGCTGGGTTGAAAGTGGCCAAACAGGTCATGTCATTTTGGAGGAGTCCCTG 1307
DB 179 GGAAGGCTGCTGGGTTGAAAGTGGCCAAACAGGTCATGTCATTTTGGAGGAGTCCCTG 238

QY 1308 GGATGGTGAAGGGAATTCAGTTTCTGAGTGGTCTGAGGAGTCCCTGCTGAGGAGTCCCTG 1367
DB 239 GGATGGTGAAGGGAATTCAGTTTCTGAGTGGTCTGAGGAGTCCCTGCTGAGGAGTCCCTG 298

QY 1368 TGGCTGGTGGTGGTGGGTCCTCCACAGTTTCTGGGTGTTCTCAGTTGGAGCAAGAGCC 1427
DB 299 TGGCTGGTGGTGGTGGGTCCTCCACAGTTTCTGGGTGTTCTCAGTTGGAGCAAGAGCC 358

QY 1428 AACTCAGGAGTGGGTCCTCCACAGCAAAATCAGAAATGAGAACTGGTAGGA 1487
DB 359 AACTCAGGAGTGGGTCCTCCACAGCAAAATCAGAAATGAGAACTGGTAGGA 418

QY 1488 GGCAGGGTGGGAGGTTGAGCTGAAGAAAGGAGGAGTTCCTGGGTCAGTTGGAGCAAGAGCC 1547
DB 419 GGCAGGGTGGGAGGTTGAGCTGAAGAAAGGAGGAGTTCCTGGGTCAGTTGGAGCAAGAGCC 478

QY 1548 CACGCTGTAATCCACAGCACTTTGGAGGTCAGAGGAGGAGTTCAGAGGTCAGAGAT 1607
DB 479 CACGCTGTAATCCACAGCACTTTGGAGGTCAGAGGAGGAGTTCAGAGGTCAGAGAT 538

QY 1608 CGAGACCATCTGCTGCTAACAAGGGTGAACCCCGTCTCTACTAAATAATCAAAAAATCA 1667
DB 1608 CGAGACCATCTGCTGCTAACAAGGGTGAACCCCGTCTCTACTAAATAATCAAAAAATCA 1667

DB 539 CGAGACCATCTGCTGCTAACAAGGTCGTGAAA--CTAATACTAATAAATAAAAAAATAA 596
QY 1668 GCCG 1671
DB 597 GCCG 600

RESULT 5
US-10-027-632-134687
; Sequence 134687, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134687
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-134687

Query Match 19.8%; Score 559.6; DB 17; Length 600;
Best Local Similarity 97.5%; Pred. No. 8e-147;
Matches 589; Conservative 1; Mismatches 10; Indels 4; Gaps 2;

QY 1068 GGCTGCTCAGCAGACGCGAAAGTGCATGTTCTGAGGAGTCTGGGGCCCTCCACGA 1127
DB 1 GGCTGCTCAGCAGACGCGAAAGTGCATGTTCTGAGGAGTCTGGGGCCCTCCACGA 60

QY 1128 CTCACGGCTCACCAGGTTTCCAGGTTAGTGGGTCCTGGCTCAGGCTGCTCATG 1187
DB 61 CTCACGGCTCACCAGGTTTCCAGGTTAGTGGGTCCTGGCTCAGGCTGCTCATG 118

QY 1188 CCCACTCTTCCCTGCTGTTGACTTCTCTGGCACCCCTCTGTCAGGCTGAGTGGGATG 1247
DB 119 CCCACTCTTCCCTGCTGTTGACTTCTCTGGCACCCCTCTGTCAGGCTGAGTGGGATG 178

QY 1248 GGAAGGCTGCTGGGTTGAAAGTGGCCAAACAGGTCATGTCATTTTGGAGGAGTCCCTG 1307
DB 179 GGAAGGCTGCTGGGTTGAAAGTGGCCAAACAGGTCATGTCATTTTGGAGGAGTCCCTG 238

QY 1308 GGATGGTGAAGGGAATTCAGTTTCTGAGTGGTCTGAGGAGTCCCTGCTGAGGAGTCCCTG 1367
DB 239 GGATGGTGAAGGGAATTCAGTTTCTGAGTGGTCTGAGGAGTCCCTGCTGAGGAGTCCCTG 298

QY 1368 TGGCTGGTGGTGGTGGGTCCTCCACAGTTTCTGGGTGTTCTCAGTTGGAGCAAGAGCC 1427
DB 299 TGGCTGGTGGTGGTGGGTCCTCCACAGTTTCTGGGTGTTCTCAGTTGGAGCAAGAGCC 358

QY 1428 AACTCAGGAGTGGGTCCTCCACAGCAAAATCAGAAATGAGAACTGGTAGGA 1487
DB 359 AACTCAGGAGTGGGTCCTCCACAGCAAAATCAGAAATGAGAACTGGTAGGA 418

QY 1488 GGCAGGGTGGGAGGTTGAGCTGAAGAAAGGAGGAGTTCCTGGGTCAGTTGGAGCAAGAGCC 1547
DB 419 GGCAGGGTGGGAGGTTGAGCTGAAGAAAGGAGGAGTTCCTGGGTCAGTTGGAGCAAGAGCC 478

QY 1548 CACGCTGTAATCCACAGCACTTTGGAGGTCAGAGGAGGAGTTCAGAGGTCAGAGAT 1607
DB 479 CACGCTGTAATCCACAGCACTTTGGAGGTCAGAGGAGGAGTTCAGAGGTCAGAGAT 538

QY 1608 CGAGACCATCTGCTGCTAACAAGGGTGAACCCCGTCTCTACTAAATAATCAAAAAATCA 1667
DB 1608 CGAGACCATCTGCTGCTAACAAGGGTGAACCCCGTCTCTACTAAATAATCAAAAAATCA 1667

```
Db 419 GGCAGGGTGGAGGGTGTGAGACTGAAGAAAGCAGAGGTTCGCGGCACGGTGGCT 478
QY 1548 CACGCTCTAATCCAGACACTTTGGAGAGCCGAGCGGGGAGATCACAGGTGAGGAGAT 1607
Db 479 CAYGCTCTAATCCAGACACTTTAGGAGGCCGAGCGGGGAGATCACAGAGGTGAGGAGAT 538
QY 1608 CGAGACCATCTGGCTAACACGGGTGAACCCCGTCTCTACTAAATAATACAAAAATCA 1667
Db 539 CGAGACCATCTGGCTAACACGGGTGAACCCCGTCTCTACTAAATAATACAAAAATCA 1667
QY 1668 GCCG 1671
Db 597 GCCG 600

RESULT 6
US-10-357-930-23318/c
; Sequence 23318, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23318
; LENGTH: 1506
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1467, 1476, 1497
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-23318

Query Match 11.9%; Score 337; DB 20; Length 1506;
Best Local Similarity 99.7%; Pred. No. 8.5e-84;
Matches 348; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 977 GGCACAGCTGTCTCTGAACGAGTTTGTGAAGGTGCCGTGCGGACAAAGTGGGTGATGAAG 1036
Db 1325 GGCACAGCTGTCTCTGAACGAGTTTGTGAAGGTGCCGTGCGGACAAAGTGGGTGATGAAG 1266
QY 1037 ATGCTGCAGATGGACATGAATCCAGCAGCTGGCTCGCTCAGCAGACGGAAGTGC 1096
Db 1265 ATGCTGCAGATGGACATGAATCCAGCAGCTGGCTCGCTCAGCAGACGGAAGTGC 1206
QY 1097 ATGTTCTGAGGAGTCTGGGGCCCTCCACGACTCCAGGCTCACCCAGGTTTCCAGGGTAG 1156
Db 1205 ATGTTCTGAGGAGTCTGGGGCCCTCCACGACTCCAGGCTCACCCAGGTTTCCAGGGTAG 1146
QY 1097 ATGTTCTGAGGAGTCTGGGGCCCTCCACGACTCCAGGCTCACCCAGGTTTCCAGGGTAG 1156
Db 1205 ATGTTCTGAGGAGTCTGGGGCCCTCCACGACTCCAGGCTCACCCAGGTTTCCAGGGTAG 1146
QY 1157 TAGGAGGGTCCCTGGCTCAGCTGCTCATGCCCACTCTTCCCTGGGTGTTGACTTTCCTG 1216
Db 1205 ATGTTCTGAGGAGTCTGGGGCCCTCCACGACTCCAGGCTCACCCAGGTTTCCAGGGTAG 1146
QY 1157 TAGGAGGGTCCCTGGCTCAGCTGCTCATGCCCACTCTTCCCTGGGTGTTGACTTTCCTG 1216
```

```
Db 1145 TAGAGGGT-CCTGGCTCAGCTGCTCATGCCCACTCTTCCCTGGGTGTTGACTTTCCTG 1087
QY 1217 GCACCCCTGTGACAGGCTGAGTGGGATGGGGAAGGCTGCTGGGTTTGAAGTGCCCAA 1276
Db 1086 GCACCCCTGTGACAGGCTGAGTGGGATGGGGAAGGCTGCTGGGTTTGAAGTGCCCAA 1027
QY 1277 CAGGTCATAGTCCATTTTGGAGGAGTCCCTGGGATGGTGAAGGGAATTC 1325
Db 1026 CAGGTCATAGTCCATTTTGGAGGAGTCCCTGGGATGGTGAAGGGAATTC 978

RESULT 7
US-10-357-930-29199/c
; Sequence 29199, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29199
; LENGTH: 1506
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1467, 1476, 1497
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-29199

Query Match 11.9%; Score 337; DB 20; Length 1506;
Best Local Similarity 99.7%; Pred. No. 8.5e-84;
Matches 348; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 977 GGCACAGCTGTCTCTGAACGAGTTTGTGAAGGTGCCGTGCGGACAAAGTGGGTGATGAAG 1036
Db 1325 GGCACAGCTGTCTCTGAACGAGTTTGTGAAGGTGCCGTGCGGACAAAGTGGGTGATGAAG 1266
QY 1037 ATGCTGCAGATGGACATGAATCCAGCAGCTGGCTCGCTCAGCAGACGGAAGTGC 1096
Db 1265 ATGCTGCAGATGGACATGAATCCAGCAGCTGGCTCGCTCAGCAGACGGAAGTGC 1206
QY 1097 ATGTTCTGAGGAGTCTGGGGCCCTCCACGACTCCAGGCTCACCCAGGTTTCCAGGGTAG 1156
Db 1205 ATGTTCTGAGGAGTCTGGGGCCCTCCACGACTCCAGGCTCACCCAGGTTTCCAGGGTAG 1146
QY 1157 TAGGAGGGTCCCTGGCTCAGCTGCTCATGCCCACTCTTCCCTGGGTGTTGACTTTCCTG 1216
Db 1145 TAGGAGGGT-CCTGGCTCAGCTGCTCATGCCCACTCTTCCCTGGGTGTTGACTTTCCTG 1087
QY 1217 GCACCCCTGTGACAGGCTGAGTGGGGAAGGCTGCTGGGTTTGAAGTGCCCAA 1276
```

```
Db 1086 GCACCCCTGTGAGGCTGAGTGGGATGGGAAGGCTGCTGGGTTTGAAGTGCCAA 1027
QY 1277 CAGGGCATAGTCCATTTTGGAGAGTCCCTCGGATCGTGAAGGAATTC 1325
Db 1026 CAGGGCATAGTCCATTTTGGAGAGTCCCTCGGATCGTGAAGGAATTC 978

RESULT 8
US-10-266-103-39
; Sequence 39, Application US/10266103
; Publication No. US20040242510A1
; GENERAL INFORMATION:
; APPLICANT: Markowitz, Sanford D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING COLON CANCERS
; FILE REFERENCE: CWRU-P01-008
; CURRENT APPLICATION NUMBER: US/10/266,103
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: 60/327537
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 62520
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-266-103-39

Query Match
Best Local Similarity 9.8%; Score 277.2; DB 20; Length 62520;
Matches 312; Conservative 0; Mismatches 38; Indels 2; Gaps 1;

QY 1518 AAAAGCAGGAGTTCGGGCGCAGGTGGCTCAGCGCTGTAATCCCAAGCACTTTGGGAGGC 1577
Db 29737 AAAATACAGATGGGCGGCGCGGTGGCTCAGCGCTGTAATCCCAAGCACTTTGGGAGGC 29796
QY 1578 CGAGCGGCGCAGATCACAGAGTTCAGAGATCGAGACCATCTGGCTAAACACGGGGTGAAA 1637
Db 29797 CGAGCGGCGCGATCACAGAGTTCAGAGATTCAGAGACCATCTGGCTAAACAC--GGTGAAA 29854
QY 1638 CCCCGTCTCTACTAAATAACAAAATACAGCCGCGGTGAGGTGGCGGCGCTGTAGTCC 1697
Db 29855 CCCCGTCTCTACTAAATAACAAAATTAGCCGGGAGGTGGCGGCGCTGTAGTCC 29914
QY 1698 CAGCTACTCAGGAGGCTGAGGCAAGAGATGGCGTGAACCCCAAGCGGCGCGAGCCTACAG 1757
Db 29915 CAGCTACTCAGGAGGCTGAGGCAAGAGATGGCGTGAACCCCAAGCGGCGCGAGCCTACAG 29974
QY 1758 TGAGCGGAGATTGGCCCACTGCACCTCAGCCTGAGCAGATGAGTCCGCTCTCAAAA 1817
Db 29975 TGAGCGGAGATTGGCCCACTGCACCTCAGCCTGAGCAGATGAGTCCGCTCTCAAAA 30034
QY 1818 AAAAAAAGAGAAAGAAAGGAGGAGTTTGGGGGCGAGGGCAGCA 1869
Db 30035 AAAAAAAGAGAAAGAAAGGAGGAGTTTGGGGGCGAGGGCAGCA 30086

RESULT 9
US-09-764-855-328/c
; Sequence 328, Application US/09764855
; Patent No. US20020119919A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA110
; CURRENT APPLICATION NUMBER: US/09/764,855
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 328
; LENGTH: 32152
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-09-764-855-328

Query Match
Best Local Similarity 9.7%; Score 275; DB 9; Length 32152;
Matches 311; Conservative 0; Mismatches 40; Indels 2; Gaps 1;

QY 1486 GAGGCAGGGTGGGAGGGTGTGAGACTGAAGAAAGGAGGAGTTGCGGGCACGGTGG 1545
Db 18284 GAGGCGGGTGGATCATGAAGTGAAGAGATCAAGACCATCTCTGTTGGCGGCGCGGTGG 18225
QY 1546 CTCAGCGCTGTAAATCCCAAGCACTTTGGGAGGCGGAGGCGGAGATCAGAGGTCAAGGAG 1605
Db 18224 CTCAGCGCTGTAAATCCCAAGCACTTTGGGAGGCGGAGGCGGCGGATCAGAGGTCAAGGAG 18165
QY 1606 ATCGAGACCATCTCGCTAAACACGGGGTGAACACCCCGTCTCTACTAAAAATACAAAAAAT 1665
Db 18164 ATCGAGACCATCTCGCTAAACAC--GGTGAACCCCGTCTCTACTAAAAATACAAAAAAT 18107
QY 1666 CAGCGGGTGAAGTGGCGGCGCCTGTAGTCCAGCTACTCAGGAGGCTGAGGCAAGAGA 1725
Db 18106 TAGCCGGGCGAGGTGGCGGCGCCTGTAGTCCAGCTACTCAGGAGGCTGAGGCAAGAGA 18047
QY 1726 ATGGCGTGAACCCCAAGGCGGCGAGCCTACAGTGAAGCGAGATTCGCCCACTGCCA 1785
Db 18046 ATGGCGTGAACCCCAAGGCGGCGAGCCTGTGCACTGAGCGAGATTCGCCCACTGCCA 17987
QY 1786 GCCTGGACGACAGTGCAGACTCCGCTCTCAAAAAAAGAAAGAAAGAAA 1838
Db 17986 GCCTGGGACAGCGAGACTCCGCTCTCAAAAAAAGAAAGAAAGAAAAGA 17934

RESULT 10
US-09-764-872-518
; Sequence 518, Application US/09764872
; Publication No. US20030050231A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA125
; CURRENT APPLICATION NUMBER: US/09/764,872
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 957
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 518
; LENGTH: 32152
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-872-518

Query Match
Best Local Similarity 9.7%; Score 275; DB 10; Length 32152;
Matches 311; Conservative 0; Mismatches 40; Indels 2; Gaps 1;

QY 1486 GAGGCAGGGTGGGAGGGTGTGAGACTGAAGAAAGGAGGAGTTGCGGGCACGGTGG 1545
Db 13869 GAGGCGGGTGGATCATGAAGTGAAGAGATCAAGACCATCTCTGTTGGCGGCGCGGTGG 13928
QY 1546 CTCAGCGCTGTAAATCCCAAGCACTTTGGGAGGCGGAGGCGGAGATCAGAGGTCAAGGAG 1605
Db 13929 CTCAGCGCTGTAAATCCCAAGCACTTTGGGAGGCGGAGGCGGCGGATCAGAGGTCAAGGAG 13988
QY 1606 ATCGAGACCATCTCGCTAAACACGGGGTGAACACCCCGTCTCTACTAAAAATACAAAAAAT 1665
Db 13989 ATCGAGACCATCTCGCTAAACAC--GGTGAACCCCGTCTCTACTAAAAATACAAAAAAT 14046
QY 1666 CAGCGGGTGAAGTGGCGGCGCCTGTAGTCCAGCTACTCAGGAGGCTGAGGCAAGAGA 1725
Db 14047 TAGCCGGGCGAGGTGGCGGCGCCTGTAGTCCAGCTACTCAGGAGGCTGAGGCAAGAGA 14106
QY 1726 ATGGCGTGAACCCCAAGGCGGCGAGCCTACAGTGAAGCGAGATTCGCCCACTGCCA 1785
```

Db 14107 ATGGCGTGAACCCAGGGGCGGAGCTGCACTGAGCCGAGATTGGCCACTGCACCTCCA 14166
QY 1786 GCCTGACGACAGTCCGCTCTCAAAAAAAGAGAAAGAAAGAA 1838
Db 14167 GCCTGGCGACGAGCACTCCGCTCTCAAAAAAAGAGAAAGAAAGAA 14219

RESULT 11
US-10-072-349-328/c
; Sequence 328, Application US/10072349
; Publication No. US20030054420A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P110C1
; CURRENT APPLICATION NUMBER: US/10/072,349
; CURRENT FILING DATE: 2002-02-11
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 328
; LENGTH: 32152
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-072-349-328

Query Match 9.7%; Score 275; DB 14; Length 32152;
Best Local Similarity 88.1%; Pred. No. 1.4e-65;
Matches 311; Conservative 0; Mismatches 40; Indels 2; Gaps 1;
QY 1486 GAGCGAGGGGTGGGAGGGTGTGAGACTGAAGAAAGGCGAGGTTGCCGGGCACGGTGG 1545
Db 18284 GAGCGGGTGGATCATGAAGTGAAGATCAAGACCATCTCTGTTGGCGGGCGGGTGG 18225
QY 1546 CTCAGCCCTGTAATCCACGACATTTGGGAGGCGGAGGCGGAGATCAAGAGTCAGAGT 1605
Db 18224 CTCAGCCCTGTAATCCACGACATTTGGGAGGCGGAGGCGGAGATCAAGAGTCAGAGT 18165
QY 1606 ATCGAGACCATCTCGCTTAACACGGGGTGAACCCCGCTCTACTAAAAATACAAAAAT 1665
Db 18164 ATCGAGACCATCTCGCTTAACAC--GGTGAACCCCGCTCTACTAAAAATACAAAAAT 18107
QY 1666 CAGCCGGTGAAGTGGCGGGCGCTGTAGTCCAGCTACTCAGAGGCTGAGGCGAAGAGA 1725
Db 18106 TAGCCGGCGAGGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAGGCTGAGGCGAAGAGA 18047
QY 1726 ATGGCGTGAACCCAGGGGCGGAGCTTACAGTGAAGCGAGATTGGCCACTGCACCTCCA 1785
Db 18046 ATGGCGTGAACCCAGGGGCGGAGCTTACAGTGAAGCGAGATTGGCCACTGCACCTCCA 17987
QY 1786 GCCTGACGACAGTCCGCTCTCAAAAAAAGAGAAAGAAAGAA 1838
Db 17986 GCCTGGCGACGAGCACTCCGCTCTCAAAAAAAGAGAAAGAAAGAA 17934

RESULT 12
US-10-087-192-1240
; Sequence 1240, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1240
; LENGTH: 129042
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1240
Query Match 9.7%; Score 275; DB 13; Length 129042;
Best Local Similarity 88.1%; Pred. No. 3.1e-65;
Matches 311; Conservative 0; Mismatches 40; Indels 2; Gaps 1;
QY 1486 GAGCGAGGGTGGGAGGGTGTGAGACTGAAGAAAGGCGAGGATTCGGCGGCACGGTGG 1545
Db 123352 GAGCGGGTGGATCATGAAGTGAAGAGATCAAGACCATCTCTGTTGGCGGGCGGGTGG 123411
QY 1546 CTCAGCCCTGTAATCCACGACATTTGGGAGGCGGAGGCGGAGATCAAGAGTCAGAGT 1605
Db 123412 CTCAGCCCTGTAATCCACGACATTTGGGAGGCGGAGGCGGAGATCAAGAGTCAGAGT 123471
QY 1606 ATCGAGACCATCTCGCTTAACACGGGGTGAACCCCGCTCTACTAAAAATACAAAAAT 1665
Db 123472 ATCGAGACCATCTCGCTTAACAC--GGTGAACCCCGCTCTACTAAAAATACAAAAAT 123529
QY 1666 CAGCGGGTGAAGTGGCGGGCGCTGTAGTCCAGCTACTCAGAGGCTGAGGCGAAGAGA 1725
Db 123530 TAGCCGGCGAGGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAGGCTGAGGCGAAGAGA 123589
QY 1726 ATGGCGTGAACCCAGGGGCGGAGCCTACAGTGAAGCGAGATTGCGCCACTGCACCTCCA 1785
Db 123590 ATGGCGTGAACCCAGGGGCGGAGCCTGAGTGAAGCGAGATTGCGCCACTGCACCTCCA 123649
QY 1786 GCCTGACGACAGTCCGCTCTCAAAAAAAGAGAAAGAAAGAA 1838
Db 123650 GCCTGGCGACAGCGAGACTCCGCTCTCAAAAAAAGAGAAAGAAAGAA 123702

RESULT 13
US-10-242-355-840/c
; Sequence 840, Application US/10242355
; Publication No. US20030235831A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC003C1
; CURRENT APPLICATION NUMBER: US/10/242,355
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,897
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 840
; LENGTH: 335
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-242-355-840

Query Match 9.7%; Score 274.6; DB 17; Length 335;
Best Local Similarity 92.0%; Pred. No. 1.5e-66;
Matches 301; Conservative 0; Mismatches 24; Indels 2; Gaps 1;

QY 1513 TGAAGAAAAAGGACAGGAGTTGCCGGGACCGGTGGCTCACGCCCTGTAAATCCAGCAGCTTTGG 1572
DB 334 TTATATAATGTCAGAGGAAGCGCGGGCGGTGGCTCACGCCCTGTAAATCCAGCAGCTTTGG 275

QY 1573 GAGGCGGAGCGGCGGAGATCACAGGTTCAGGAGATCAGAGACCATCTGGCTTAACACGGGG 1632
DB 274 GAGGCGGAGCGGCTGGATCACAGGTTCAGGAGATCAGAGACCATCTGGCTTAACAC--GG 217

QY 1633 TGAACCCCGCTCTCTACTATAAATAACAAAAATCAGCCGGGTGAGGTGGCGGGCGGCTGT 1692
DB 216 TGAACCCCGCTCTCTACTATAAATAACAAAAATTAAGCCGGGAGGTGGCGGGCGGCTGT 157

QY 1693 AGTCCAGCTACTCAGGAGGCTGAGGCAAGAGAAATGGCGTGAACCCCGGCGGCGGAGCC 1752
DB 156 AGTCCAGCTACTCAGGAGGCTGAGGCAAGAGAAATGGCGTGAACCCCGGCGGCGGAGCC 97

QY 1753 TACAGTGAGCGAGATTTGCCCACTGCACTCCAGCCTGGACGACAGTGAAGACTCCGCTCTC 1812
DB 96 TCGAGTGAGCGAGATTTGCCCACTGCACTCCAGCCTGGGCGACAGAGACTCCGCTCTC 37

QY 1813 AAAAAAAGAAAAAGAAAAA 1839
DB 36 AAAAAAAGAAAAAAGAAAAA 10

RESULT 14
US-10-087-192-544
; Sequence 544, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 544
; LENGTH: 238484
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-544

Query Match 9.7%; Score 273.4; DB 13; Length 238484;
Best Local Similarity 90.2%; Pred. No. 1.2e-64;
Matches 304; Conservative 0; Mismatches 31; Indels 2; Gaps 1;

QY 1503 GTGTTGAGACTCAAGAAAGGACAGGATTTGCCGGGACCGGTGGCTCACGCCCTGTAAATCC 1562
DB 38352 GTGGCAAAAGTGTGAAAGAAATGATGGCGCGGGCGGTGGCTCACGCCCTGTAAATCC 38411

QY 1563 AGCATTTCGGAGGCGGCGGAGATCACAGGTTCAGGAGATCGAGACCATCTTGGC 1622
DB 38412 AGCATTTCGGAGGCGGCGGAGATCACAGGTTCAGGAGATCGAGACCATCTTGGC 38471

QY 1623 TAAACGGGTGAACCCCGTCTCTACTATAAATAACAAAAATCAGCCGGGTGAGGTGGC 1682
DB 38472 TAAAC--GGTGAACCCCTGTCTCTACTATAAATAACAAAAATTAAGCCGGGCGGAGTGGC 38529

QY 1683 GGGCGCTGTAGTCCAGCTACTCAGGAGGTGAGGCAAGAGAAATGGCGTGAACCCAGG 1742
DB 38530 GGGCGCTGTAGTCCAGCTACTCAGGAGGTGAGGCAAGAGAAATGGCGTGAACCCAGG 38589

QY 1743 GGGCCGAGCCTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 1802
DB 38590 GGGCCGAGCCTGTGAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 38649

QY 1803 ACTCCGTCTCAAAAAAAGAAAAAAGAAAAAAGAAAAA 1839
DB 38650 ACTCCGTCTCAAAAAAAGAAAAAAGAAAAAAGAAAAAATA 38686

RESULT 15
US-10-220-510-5/c
; Sequence 5, Application US/10220510
; Publication No. US20030190637A1
; GENERAL INFORMATION:
; APPLICANT: Hovnanian, Alain
; APPLICANT: Chavanas, Stephanie
; APPLICANT: Cookson, William
; APPLICANT: Moffat, Miriam
; APPLICANT: Walley, Andrew
; TITLE OF INVENTION: SUSCEPTIBILITY GENE FOR NETHERTON'S DISEASE
; FILE REFERENCE: I00317.70008.US
; CURRENT APPLICATION NUMBER: US/10/220,510
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: GB 0005098.9
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: GB 0005229.0
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7995
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7994)..(7995)
; OTHER INFORMATION: n = a, c, g or t/u
US-10-220-510-5

Query Match 9.7%; Score 272.6; DB 16; Length 7995;
Best Local Similarity 94.8%; Pred. No. 3.2e-65;
Matches 293; Conservative 0; Mismatches 14; Indels 2; Gaps 1;

QY 1532 GCCGGGACCGGTGGCTCACGCCCTGTAAATCCAGCAGCTTTGGAGGCGCGGCGGAGAT 1591
DB 3681 GCCGGGCGGTGGCTCACGCCCTGTAAATCCAGCAGCTTTGGAGGCGCGGCGGAGAT 3622

QY 1592 CACGAGGTTCAGGAGATCAGAGACCATCTGGCTTAACACGGGGTGAACCCCGTCTCTACTA 1651
DB 3621 CACGAGGTTCAGGAGATCAGAGACCATCTGGCTTAACAC--GGTGAACCCCGTCTCTACTA 3564

QY 1652 AAATACAAAAAATCAGCGGTGAGGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAG 1711
DB 3563 AAATACAAAAAATTAGCGGCGGAGGTGGCGGGCGCTGTAGTCCAGCTACTCAGGAG 3504

QY 1712 GCTGAGGCAAGAGAAATGGCGTGAACCCAGGCGGCGGAGCCTTACAGTACGCGGAGATTGC 1771
DB 3503 GCTGAGGCAAGAGAAATGGCGTGAACCCAGGCGGCGGAGCCTTACAGTACGCGGAGATTGC 3444

QY 1772 GCCACTGCATCCAGCTTCGAGCAGCAGTGAAGTCCGCTCTCAAAAAAAGAAAAA 1831
DB 3443 GCCACTGCATCCAGCTTCGAGCAGCAGTGAAGTCCGCTCTCAAAAAAAGAAAAA 3384

QY 1832 AAAGAAAAAG 1840
DB 3383 AAAAAAAG 3375

Search completed: June 29, 2005, 12:01:59
Job time : 1107.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2005, 15:06:07 ; Search time 6070.87 Seconds
(without alignments)
17687.639 Million cell updates/sec

Title: US-08-731-499-6
Perfect score: 2821
Sequence: 1 ATCTTAGCGCACACGCTG.....AGGGGTTACAGACGACCTG 2821

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hgc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	766	27.2	779	5	BX111942	BX111942 BX111942
2	615.4	21.8	668	4	BM690083	BM690083 UI-E-CK1-
3	560.8	19.9	572	4	BM690003	BM690003 UI-E-CK1-
4	555	19.7	632	5	BQ637863	BQ637863 hdl5e08.y
5	537	19.0	559	4	BM662840	BM662840 UI-E-CK0-
6	524.6	18.6	546	5	BQ447506	BQ447506 UI-H-EU1-
7	520.8	18.5	672	2	AW967092	AW967092 EST379166
8	500.2	17.7	534	4	BM661803	BM661803 UI-E-CK1-
9	486	17.2	486	2	BE220617	BE220617 ht99c12.x
10	466.8	16.5	471	1	AI377509	AI377509 tc37d08.x
11	464.8	16.5	489	4	BM688117	BM688117 UI-E-CL0-
12	461.8	16.4	489	1	AI424326	AI424326 te90c12.x
13	456.4	16.2	475	4	BM704154	BM704154 UI-E-CK1-
14	448.8	15.9	455	2	AW020233	AW020233 wt83d08.x
15	436	15.5	437	4	BM686135	BM686135 UI-E-CK0-
16	418.2	14.8	439	1	AI306391	AI306391 qn12f08.x
17	408.6	14.5	442	4	BM672390	BM672390 UI-E-CL1-
18	397.2	14.1	407	4	BM701070	BM701070 UI-E-CL1-
19	380.8	13.5	399	4	BI293114	BI293114 UI-R-DK0-
20	376.4	13.3	511	8	AQ767124	AQ767124 HS 2206.B
21	374	13.3	407	7	H40682	H40682 yp50g05.ai
22	363	12.9	378	2	AW136419	AW136419 UI-H-B11-
23	335.8	11.9	404	1	AA308846	AA308846 EST179652
24	325	11.5	740	7	W27113	W27113 20d8 Human

C	25	324.8	11.5	456	7	W26445	W26445 29f7 Human
	26	321.8	11.4	608	7	W27002	W27002 19h12 Human
	27	319	11.3	782	7	W28792	W28792 52f3 Human
C	28	314.4	11.1	579	7	W27379	W27379 28c1 Human
	29	311.8	11.1	331	1	AA317761	AA317761 EST19742
C	30	311	11.0	829	7	W26485	W26485 30c1 Human
	31	302.6	10.7	483	7	W27135	W27135 22h4 Human
C	32	296.4	10.5	843	7	W28227	W28227 43h1 Human
	33	293.4	10.4	342	1	AA317311	AA317311 EST19308
	34	276	9.8	760	7	W27109	W27109 20d3 Human
	35	267.6	9.5	532	8	AQ184326	AQ184326 HS 3220.A
	36	266	9.4	267	1	AA378172	AA378172 EST90835
	37	264.8	9.4	5670	3	HS804253	AL832942 Homo sapi
C	38	264.4	9.4	311	5	BQ429229	BQ429229 AGENCOURT
	39	260.2	9.2	295	5	BUT727375	BUT727375 UI-E-CQ1-
	40	259.6	9.2	498	8	AQ199645	AQ199645 RPI111-62
	41	258.6	9.2	570	7	CR542495	CR542495 DKF2p459J
	42	258.6	9.2	630	7	CR542494	CR542494 DKF2p459J
	43	256.4	9.1	409	8	AQ627863	AQ627863 CTBT1-E1-
C	44	256.2	9.1	334	5	BX493558	BX493558 DKF2p781M
C	45	252.2	8.9	297	4	BM663617	BM663617 UI-E-CK0-

ALIGNMENTS

RESULT 1
BX111942
LOCUS BX111942 779 bp mRNA linear EST 07-FEB-2003
DEFINITION BX111942 Soares retina N2b4HR Homo sapiens cDNA clone
IMAGP998H09360 ; IMAGE:190904, mRNA sequence.
ACCESSION BX111942
VERSION BX111942.1 GI:27878691
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 779)
AUTHORS Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
Radelof, U., Schneider, D. and Korn, B.
TITLE Human Unigeneset - RZPD3
JOURNAL Unpublished (2003)
COMMENT Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD: IMAGP998H09360
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human Unigeneset - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi?response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTCACAGGAAACAGCTATGAC.

FEATURES
source

1..779
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGP998H09360 ; IMAGE:190904"
/sex="male"
/tissue_type="retina"
/dev_stage="55 year old"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares retina N2b4HR"
/note="Organ: eye; Vector: pT7F3D (Pharmacia) with a
modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st
strand cDNA was primed with a Not 1 - oligo(dT) primer [5'

TGTTACCAATCTGAAGTGGAGCGCGCGCTTTTTTTTTTTTTTTTTTTT 3'}.
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). The retinas were obtained from a 55 year old
Caucasian and total cellular poly(A)+ RNA was extracted 6
hrs after their removal. The retina RNA was kindly
provided by Roderick R. McInnes M.D. Ph.D. from the
University of Toronto. Library constructed by Bento
Soares and M.Fatima Bonaldo. "

ORIGIN		Query Match	27.2%;	Score 766;	DB 5;	Length 779;
		Best Local Similarity	99.6%;	Pred. No. 2.3e-126;		
		Matches 777;	Conservative 0;	Mismatches 2;	Indels 1;	Gaps 1;
QY	1577	CCGAGCGCGGCGAGATCAGAGTTCAGGAGATCGAGACCAATCTTGCTTAACAGCGGGTGAA	1636			
Db	1	CCGAGCGCGGCGAGAT-NCGAGGTGAGGAGATCGAGACCAATCTTGCTTAACAGCGGGTGAA	59			
QY	1637	ACCCGCTCTCTACTAAATATACAAATATCAGCGGGTGAGTGGGGCGCTGTAGTC	1696			
Db	60	ACCCGCTCTCTACTAAATATACAAATATCAGCGGGTGAGTGGGGCGCTGTAGTC	119			
QY	1697	CCAGCTACTCAGAGGCTGAGGCAAGAGAAATGGCGTGAACCCAGCGGGCGGAGCTTACA	1756			
Db	120	CCAGCTACTCAGAGGCTGAGGCAAGAGAAATGGCGTGAACCCAGCGGGCGGAGCTTACA	179			
QY	1757	GTGAGCGGAGATTTGGCCCATCTGCACTCCAGCTTGACGACAGTGCAGTCCGTCTCAAAA	1816			
Db	180	GTGAGCGGAGATTTGGCCCATCTGCACTCCAGCTTGACGACAGTGCAGTCCGTCTCAAAA	239			
QY	1817	AAAAAAAAGAAAGAAAGAAAGGAGAGATTTGGGGGGGAGGGGAGCAATATTC	1876			
Db	240	AAAAAAAAGAAAGAAAGAAAGGAGAGATTTGGGGGGGAGGGGAGCAATATTC	299			
QY	1877	TATAACTTCGGGATGCTGAGGGGGCTTCATGGGAGGAGCCCTGGCTCTCTCCCAAA	1936			
Db	300	TATAACTTCGGGATGCTGAGGGGGCTTCATGGGAGGAGCCCTGGCTCTCTCCCAAA	359			
QY	1937	GGCATCTCTCACAGTGGTGTCAACAGGAAAAATGGCAGCAAAATACGCTGAGGCTGTGT	1996			
Db	360	GGCATCTCTCACAGTGGTGTCAACAGGAAAAATGGCAGCAAAATACGCTGAGGCTGTGT	419			
QY	1997	CTTTCTGCTTTGAAAGGTCAGCTGTACTTAAAGGAGTGTTCAGCTCTGCTGGGTG	2056			
Db	420	CTTTCTGCTTTGAAAGGTCAGCTGTACTTAAAGGAGTGTTCAGCTCTGCTGGGTG	479			
QY	2057	CTGCTCTGGGACCCCTGCTGCAACCCACCACTCCCCCAACATCTCTCTTCCATCC	2116			
Db	480	CTGCTCTGGGACCCCTGCTGCAACCCACCACTCCCCCAACATCTCTCTTCCATCC	539			
QY	2117	ATATCCCCCAGTATGAGACCTTCCACAACTCCAGCCATAAGCTGAATGTTCTTTTAA	2176			
Db	540	ATATCCCCCAGTATGAGACCTTCCACAACTCCAGCCATAAGCTGAATGTTCTTTTAA	599			
QY	2177	GGATGAGAAACTTCTGTCTGTCTGCAAGAAATGGGGAGCTGTGACTGGGATGT	2236			
Db	600	GGATGAGAAACTTCTGTCTGTCTGTCTGCAAGAAATGGGGAGCTGTGACTGGGATGT	659			
QY	2237	GGGCTGGGCTTGGCTTCTAACTGTGTGTGACCCAGAGCAGTCTTCTCTCCCTAAC	2296			
Db	660	GGGCTGGGCTTGGCTTCTAACTGTGTGTGACCCAGAGCAGTCTTCTCTCCCTAAC	719			
QY	2297	TTGGTTATGTCTTGGCAGCAGTGAAGAGGTCGGA CTAGGGCGAAACAGTTTTGGATTAT	2356			
Db	720	TTGGTTATGTCTTGGCAGCAGTGAAGAGGTCGGA CTAGGGCGAAACAGTTTTGGATTAT	779			

RESULT 2
BM690083
LOCUS
DEFINITION UI-E-CK1-abr-f-10-0-UI.r1 UI-E-CK1 Homo sapiens cDNA clone

UI-E-CK1-abr-f-10-0-UI 5', mRNA sequence.
BM690083
Version 1 GI:19003341
EST.
Keywords
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 668)
Bonaldo,M.F., Lemmon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
MEDLINE 97044477
PubMed 8889548
COMMENT
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.
Location/Qualifiers
1..668
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-CK1-abr-f-10-0-UI"
/tissue_type="Retina Foveal and Macular"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-CK1"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-CK1 is a normalized cDNA library containing the
following tissue(s): Retina Foveal and Macular. The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT73-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is GTCC. This library was created for the program,
Gene Discovery in the Visual System, supported by National
Eye Institute (NEI)."

ORIGIN

Query Match		21.8%;	Score 615.4;	DB 4;	Length 668;
		Best Local Similarity	97.0%;	Pred. No. 1.6e-99;	
		Matches 650;	Conservative 0;	Mismatches 12;	Indels 8; Gaps 2;
QY	1861	GGGGCAGCAATATTTCTATTAACCTCCGGGATGCTGAGGGGCGTTTCATGGGAGGACCCCTG	1920		
Db	7	GGGGCAGCAATATTTCTATTAACCTCCGGGATGCTGAGGGGCGTTTCATGGGAGGACCCCTG	66		
QY	1921	GCCTCTCTCCCCAAGGCATCTCACCAGTGGTGTCAACAGAAAAATGGCAGCAATA	1980		
Db	67	GCCTCTCTCCCCAAGGCATCTCACCAGTGGTGTCAACAGAAAAATGGCAGCAATA	126		
QY	1981	CGCTCAGGCTGTGCTCTTCTGCGCTTTGAAAGGGTTCAGCTGACTTAAGGAGCTGTTT	2040		
Db	127	CGCTCAGGCTGTGCTCTTCTGCGCTTTGAAAGGGTTCAGCTGACTTAAGGAGCTGTTT	186		

tissue(s): Retina Foveal and Macular. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GTCC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).

TAG_TISSUE=Foveal and Macular Retina
TAG_LIB=UI-B-CK0
TAG_SEQ=GTCC"

ORIGIN

Query Match 19.0%; Score 537; DB 4; Length 559;
Best Local Similarity 99.1%; Pred. No. 1.6e-85;
Matches 540; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2095 CAACATCTCTCTTCCATCCATATCCCCAGTATGGACCTTCCACAACTCCAGCCAT 2154
DB 559 CAACATCTCTCTTCCATCCATATCCCCAGTATGGACCTTCCACAACTCCAGCCAT 500

QY 2155 AAGCTGAATGTTTCTCTTTAAAGGATGGAGAAACTTCTGTCTCTGGCAAGAAATG 2214
DB 499 AAGCTGAATGTTTCTCTTTAAAGGATGGAGAAACTTCTGTCTCTGGCAAGAAATG 440

QY 2215 GGGGACTGTGTGCTGGATTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 2274
DB 439 GGGGACTGTGTGCTGGATTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 380

QY 2275 CAGCCACTTCTCTCTTAACTTGTATGTTGGCAGCAGTGGAGCAGTGGAGT 2334
DB 379 CAGCCACTTCTCTCTTAACTTGTATGTTGGCAGCAGTGGAGCAGTGGAGT 320

QY 2335 AGCGCAACAGTTTGGATTTGTTTATAGTGTGGAATTTATTTTGTATATAAAC 2394
DB 319 AGCGCAACAGTTTGGATTTGTTTATAGTGTGGAATTTATTTTGTATATAAAC 260

QY 2395 TCTTATGTATACCCCAATATAGAACTAGATTTAAAGGAGTCTCTCTGGTTGAAAGG 2454
DB 259 TCTTATGTATACCCCAATATAGAACTAGATTTAAAGGAGTCTCTCTGGTTGAAAGG 200

QY 2455 GAGCTGAGTACCTCTGGAACTGGAGGACCTCTGAAAGCAAACTGAAACCAAGTGC 2514
DB 199 GAGCTGAGTACCTCTGGAACTGGAGGACCTCTGAAAGCAAACTGAAACCAAGTGC 140

QY 2515 CTTGGGTCACTGTTACTCTTAAAGACAGTGTAAAGTGGAGCTGGAAAAACATTTGCTT 2574
DB 139 CTTGGGTCACTGTTACTCTTAAAGACAGTGTAAAGTGGAGCTGGAAAAACATTTGCTT 80

QY 2575 TACCTTGAATAGATAGTTTATGTTGGTATATAGAAATATAACCTATTAAACC 2634
DB 79 TACCTTGAATAGATAGTTTATGTTGGTATATAGAAATATAACCTATTAAACC 20

QY 2635 TGAGA 2639
DB 19 TGAAA 15

RESULT 6

BO447506/c
LOCUS BO447506 546 bp mRNA linear EST 29-MAY-2002
DEFINITION UI-H-EUI-bae-m-11-0-UI.s1 NCI CGAP Ctl Homo sapiens cDNA clone
UI-H-EUI-bae-m-11-0-UI 3', mRNA sequence.
ACCESSION BO447506
VERSION BO447506.1 GI:21250611
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS
JOURNAL
COMMENT

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 546)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@iowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source

Location/Qualifiers
1..546
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-EUI-bae-m-11-0-UI"
/tissue_type="Osteoarthritic Cartilage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP Ctl"
/note="Organ: Knee; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Ctl is a normalized cDNA library containing the following tissue(s): Osteoarthritic Cartilage The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TGATCACGCT.
TAG TISSUE=osteoarthritic cartilage
TAG_LIB=UI-H-EUI
TAG_SEQ=TGATCACGCT"

ORIGIN

Query Match 18.6%; Score 524.6; DB 5; Length 546;
Best Local Similarity 99.2%; Pred. No. 2.6e-83;
Matches 527; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2109 TTCCATCCATATCCCCAGTATGGACCTTCCACAACTCCAGCCATAAGCTGAATGTTTC 2168
DB 546 TTCCATCCATATCCCCAGTATGGACCTTCCACAACTCCAGCCATAAGCTGAATGTTTC 487

QY 2169 TCTTTAAGATGGAGAAAACTTCTGTCTCTCTGGCAAGAAATTTGGGGAGCTGTTGACT 2228
DB 486 TCTTTAAGATGGAGAAAACTTCTGTCTCTCTGGCAAGAAATTTGGGGAGCTGTTGACT 427

QY 2229 GGGATTTGGGCTGGCTTGGCTTCTTAAGTGTGTGACCCCAAGACACCCACTTCTCCT 2288
DB 426 GGGATTTGGGCTGGGCTTGGCTTCTTAAGTGTGTGACCCCAAGACACCCACTTCTCCT 367

QY 2289 CCCTAACCTTGGTTATGTCTTGGCAGCAGTGGAGCTCGGACTAGCGCAACAGTTTTT 2348
DB 366 CCCTAACCTTGGTTATGTCTTGGCAGCAGTGGAGCTCGGACTAGCGCAACAGTTTTT 307

QY 2349 GGATTTATGTTTTAGATGGAAATTTATTTTGTATATAAACTTTATGTGTAAACC 2408
DB 306 GGATTTATGTTTTAGATGGAAATTTATTTTGTATATAAACTTTATGTGTAAACC 247

QY 2409 CCATATAGAACTAGATTTAAAGGAGTCTCTCTGGTTGAAAGGGAGCTAGTACCTT 2468
DB 246 CCATATAGAACTAGATTTAAAGGAGTCTCTCTGGTTGAAAGGGAGCTAGTACCTT 187

```

QY 2469 CTGGAACCTGGAGCACCTCTCTGAAAAAGCAAACTGAAACAGTGCCCTGGGTCACTGTT 2528
Db 186 CTGGAACTCTGGAGCACCTCTCTGAAAAAGCAAACTGAAATCTAGTGCCTGGGTCACTGTT 127
QY 2529 ACTCCTATAAGACAGTCTTAAAGTGGAGACCTGGAAAAACAATTTGCTTTTACCTTTGAATAGAT 2588
Db 126 ACTCCTATAAGACAGTCTTAAAGTGGAGACCTGGAAAAACAATTTGCTTTTACCTTTGAATAGAT 67
QY 2589 AGGTTTTATGTTGTTATATAAGAAATAAACTAAACCTATTAAACCTGAGA 2639
Db 66 AGGTTTTATGTTGTTATATAAGAAATAAACTAAACCTATTAAACCTGAAA 16

RESULT 7
AW967092 672 bp mRNA linear EST 01-JUN-2000
LOCUS EST379166 MAGe resequences, MAGJ Homo sapiens cDNA, mRNA sequence.
ACCESSION AW967092
VERSION AW967092.1 GI:8156928
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 672)
AUTHORS Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C.,
Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
Quackenbush,J.
TITLE Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
Plate: 236
Seq primer: Reverse.
FEATURES
    source
    1..672
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone_lib="MAGe resequences, MAGJ"
    /note="Vector: pBluescriptSKm"

ORIGIN
Query Match 18.5%; Score 520.8; DB 2; Length 672;
Best Local Similarity 99.4%; Pred. No. 1.2e-82; Indels 1; Gaps 1;
Matches 533; Conservative 0; Mismatches 2;

QY 2091 CCCCCAACAAATCCTCTCTTTCCATCCATATCCCCAGTAGTGACCTTCACAACTCCCGAG 2150
Db 1 CCCCCAACAAATCCTCTCTTTCCATCCATATCCCCAGTAGTGACCTTCACAACTCCCGAG 60
QY 2151 CCATAAGCTGAATGTTTCTCTTTAAAGGATGAGAAAACTTCTGTCTGTCTCTGCGAAGA 2210
Db 61 CCATAAGCTGAATGTTTCTCTTTAAAGGATGAGAAAACTTCTGTCTGTCTGCGAAGA 120
QY 2211 ATTGGGGGACTGTGTACTGGGATTTGGGCTGGGCTTGTCTTAAGTCTGTGTGACCC 2270
Db 121 ATTGGGGGACTGTGTACTGGGATTTGGGCTGGGCTTGTCTTAAGTCTGTGTGACCC 180
QY 2271 AAGACAGCCACTTCTCCTCCCTAAACCTTGGTTATGTCTTTGGCAGCAGTGGAGGTCG 2330
Db 181 AAGACAGCCACTTCTCCTCCCTAAACCTTGGTTATGTCTTTGGCAGCAGTGGAGGTCG 240
QY 2331 GACTAGGCGAACAGTTTGGATATATGCTTTTGTAGATGGAATATTTTGTATAT 2390
Db 241 GACTAGGCGAACAGTTTGGATATATGCTTTTGTAGATGGAATATTTTGTATAT 300

```

```

QY 2391 AAACCTTTATGTGAACCCCAATATAGAAACTAGATTTAAAGGGAGTCTCTCTGGTTGAA 2450
Db 301 AAACCTTTATGTGAACCCCAATATAGAAACTAGATTTAAAGGGAGTCTCTCTGGTTGAA 360
QY 2451 AGGGAGCTGAGTACCCCTCTGGAACCTGGAGGACCTCTGAAAAAGCAAACTGAAAAACA 2510
Db 361 AGGGAGCTGAGTACCCCTCTGGAACCTGGAGGACCTCTGAAAAAGCAAACTGAAAAACA 420
QY 2511 GTGCCCTGGGTACCTGTTACTCTTATAGACAGATTTTAAAGTGAGACCTGGAACAACTTT 2570
Db 421 GTGCCCTGGGTACCTGTTACTCTTATAGACAGATTTTAAAGTGAGACCTGGAACAACTTT 480
QY 2571 GCTTTACCTTGAATAGATAGGTTTTTATCTTGGTATAT-AAGAAATAAACTAAACC 2625
Db 481 GCTTTACCTTGAATAGATAGGTTTTTATCTTGGGATATAAGAAATAAACTAAACC 536

RESULT 8
BM661803 534 bp mRNA linear EST 27-FEB-2002
LOCUS UI-E-CK1-abr-f-10-0-UI.s1 UI-E-CK1 Homo sapiens cDNA clone
DEFINITION UI-E-CK1-abr-f-10-0-UI 3', mRNA sequence.
ACCESSION BM661803
VERSION BM661803.1 GI:18965842
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 534)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Forward
POLYA=Yes.
FEATURES
    source
    1..534
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="UI-E-CK1-abr-f-10-0-UI"
    /tissue_type="Retina Foveal and Macular"
    /dev_stages="adult"
    /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
    /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
    modified polylinker; Site 1: EcoR I; Site 2: Not I;
    UI-E-CK1 is a normalized cDNA library containing the
    following tissue(s): Retina Foveal and Macular. The
    library was constructed according to Bonaldo, Lennon and
    Soares, Genome Research, 6:791-806, 1996. First strand
    cDNA synthesis was primed with an oligo-dT primer
    containing a Not I site. Double stranded cDNA was ligated
    to an EcoR I adaptor, digested with Not I, and cloned
    directionally into pT73-Pac vector. The oligonucleotide
    used to prime the synthesis of first-strand cDNA contains
    a library tag sequence that is located between the Not I

```

site and the (dT)18 tail. The sequence tag for this library is GRCC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
TAG TISSUE=Poveal and Macular Retina
TAG_LIB=UI-E-CK1
TAG_SEQ=GTCC"

ORIGIN

Query Match 17.7%; Score 500.2; DB 4; Length 534;
Best Local Similarity 98.7%; Pred. No. 5.9e-79;
Matches 516; Conservative 0; Mismatches 3; Indels 4; Gaps 1;
2117 ATATCCCCAGTAGTGACCTTCACAACTCCAGCCATAGCTGAATGTTCTCTTTAA 2176
Db |||||
534 ATATCCCCAGTAGTGACCTTCACAACTCCAGCCATAGCTGAATGTTCTCTTTAA 475
Qy |||||
2177 GGATGAGAGAAACTTCTGCTGCTCTGCGCAAGATTGGGGACTCTGACCTGGGATTCT 2236
Db |||||
474 GGATGAGAGAAACTCTGCTGCTCTGCGCAAGATTGGGGACTCTGACCTGGGATTCT 419
Qy |||||
2237 GGCTGGGCTTGCTTCTAACTCTGCTGACCAAGACAGCCACTTCTCTCCCTAAC 2296
Db |||||
418 GGCTGGGCTTGCTTCTAACTCTGCTGACCAAGACAGCCACTTCTCTCCCTAAC 359
Qy |||||
2297 TTGGTTATGCTTTGGCAGCAGCAGTAGGAGGTCGGAAGAGGAGGAGGAGGAGGAGG 2356
Db |||||
358 TTGGTTATGCTTTGGCAGCAGCAGTAGGAGGTCGGAAGAGGAGGAGGAGGAGGAGG 299
Qy |||||
2357 GTGTTTTAGTGTGAATATTTTGTATATAACTCTTATGTAACCCCAATATA 2416
Db |||||
298 GTGTTTTAGTGTGAATATTTTGTATATAACTCTTATGTAACCCCAATATA 239
Qy |||||
2417 GAACTAGATTAAGAGGAGTCTCTGTTGAAAGGGAGCTGAGTACCTCTGGA 2476
Db |||||
238 GAACTAGATTAAGAGGAGTCTCTGTTGAAAGGGAGCTGAGTACCTCTGGA 179
Qy |||||
2477 GGAGGCACCTTGAAAAAGCAAACTGAAACAGTGCCTGGTCACTGTTACTCTAT 2536
Db |||||
178 GGAGGCACCTTGAAAAAGCAAACTGAAACAGTGCCTGGTCACTGTTACTCTAT 119
Qy |||||
2537 AAGACAGTTTAAGTGAAGCTGGAAACATTTGCTTTACCTTGAATAGATGTTT 2596
Db |||||
118 AAGACAGTTTAAGTGAAGCTGGAAACATTTGCTTTACCTTGAATAGATGTTT 59
Qy |||||
2597 ATGTTGGTATTAAGAAATAAACTAACTATTAACCTGAGA 2639
Db |||||
58 ATGTTGGTATTAAGAAATAAACTAACTATTAACCTGAGA 16

RESULT 9
BE220617/c
LOCUS
DEFINITION
ht99c12.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3165238 3',
mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 486)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL, send email to:
info@image.llnl.gov

Seq primer: -40UP from Gibco

High quality sequence stop: 450.

FEATURES
source

Location/Qualifiers
1..486
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3165238"
/tissue_type="carcinoid"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu24"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; plasmid DNA from the normalized
library NCI_CGAP_Lu5 was prepared, and ss circles were
made in vitro. Following HAP purification, this DNA was
used as tracer in a subtractive hybridization reaction.
The driver was PCR-amplified cDNAs from a pool of 5,000
clones made from the same library (clones IDs
1414920-1417991 and 1520904-1522439). Subtraction by Bento
Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 17.2%; Score 486; DB 2; Length 486;
Best Local Similarity 100.0%; Pred. No. 2e-76;
Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
2143 ACTCCAGCATAAGCTGAATGTTCTCTTTAAAGATGGAGAAACTTCTGCTGCTC 2202
Db |||||
486 ACTCCAGCATAAGCTGAATGTTCTCTTTAAAGATGGAGAAACTTCTGCTGCTC 427
Qy |||||
2203 TGSCAAGATTGGGGACCTGTTGACCTGGGATTTGCTGGGCTGGCTTCTTAACCTG 2262
Db |||||
426 TGSCAAGATTGGGGACCTGTTGACCTGGGATTTGCTGGGCTGGCTTCTTAACCTG 367
Qy |||||
2263 TGTGACCCAGACAGCAGCTTCTCTCCCTAACCTTGGTATGCTTTGGCAGCAGTGA 2322
Db |||||
366 TGTGACCCAGACAGCAGCTTCTCTCCCTAACCTTGGTATGCTTTGGCAGCAGTGA 307
Qy |||||
2323 GCAGGTCGACATAGGCGAACAGTTTGGATTTATGTTTTAGATGTGAATTTT 2382
Db |||||
306 GCAGGTCGACATAGGCGAACAGTTTGGATTTATGTTTTAGATGTGAATTTT 247
Qy |||||
2383 TGTATATAAACTTTATGTTAAACCCCAATATAGAACTAGATTAAAGGGAGTCTCTC 2442
Db |||||
246 TGTATATAAACTTTATGTTAAACCCCAATATAGAACTAGATTAAAGGGAGTCTCTC 187
Qy |||||
2443 TGTGTAAGGGGAGCTGAGTACCTCTCGAACTGGAGGACCTCTGAAAAAGCAAACT 2502
Db |||||
186 TGTGTAAGGGGAGCTGAGTACCTCTCGAACTGGAGGACCTCTGAAAAAGCAAACT 127
Qy |||||
2503 GAAACACAGTGCCTGGGTCACCTGTTAATCTCTATTAAGACAGTTTAAAGTGAAG 2562
Db |||||
126 GAAACACAGTGCCTGGGTCACCTGTTAATCTCTATTAAGACAGTTTAAAGTGAAG 67
Qy |||||
2563 AAACATTTCTTTACCTTGAATAGATGCTTTTATGTTGGTATATAAGAAATAAACTA 2622
Db |||||
66 AAACATTTCTTTACCTTGAATAGATGCTTTTATGTTGGTATATAAGAAATAAACTA 7
Qy |||||
2623 ACCTAT 2628
Db |||||
6 ACCTAT 1

RESULT 10
AI377509/c
LOCUS
DEFINITION
AI377509 471 bp mRNA linear EST 18-MAR-1999
tc37d08.x1 Soares total fetus_Nb2HF8_9w Homo sapiens cDNA clone
IMAGE:2066799 3' similar to contains element MER22 repetitive
element i, mRNA sequence.

ACCESSION AI377509
 VERSION AI377509.1 GI:4187362
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 471)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapsb@mail.nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 816 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 428.
 Location/Qualifiers
 FEATURES
 source
 1..471
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2066799"
 /dev_stage="8-9 weeks"
 /lab_host="DH10B"
 /clone_lib="Soares total_fetus Nb2HF8_9w"
 /note="Vector: pT730-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from pooled 8-9 week
 (total) fetus material with a Not I - oligo(dT) primer [5'
 TGTTACCAATCTGAAGTGGAGCGCCCTTAATTTTCTTTTCTTTT 3'].
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaudo. "
 ORIGIN
 Query Match 16.5%; Score 466.8; DB 1; Length 471;
 Best Local Similarity 99.4%; Pred. No. 5.4e-73;
 Matches 468; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2169 TCTTTAAAGGATGGAGAAAACCTTCTGCTCTCTGGCAAGAAATGGGGAGCTGTGACT 2228
 DB 471 TCTTTAAGATGGAGAAACCTTCTGCTCTCTGGCAAGAAATGGGGAGCTGTGACT 412
 QY 2229 GGGATTGTGGGCTGGGCTTGGCTTCTTAAGTGTGTGACCAAGACAGCCACTTCTCT 2288
 DB 411 GGGATTGTGGGCTGGGCTTGGCTTCTTAAGTGTGTGACCAAGACAGCCACTTCTCT 352
 QY 2289 CCCTAACCTTGTTATGCTTGGGACGACAGTGAGGAGTGGAGCTAGCGAACAGTTT 2348
 DB 351 CCCTAACCTTGTTATGCTTGGGACGACAGTGAGGAGTGGAGCTAGCGAACAGTTT 292
 QY 2349 GGATTATTGTGTTTTAGATGGAAATATTTTTTGTATATAAACTTATGTGTAACC 2408
 DB 291 GGATTATTGTGTTTTAGATGGAAATATTTTTTGTATATAAACTTATGTGTAACC 232
 QY 2409 CCAATATGAAACTAGATTAAAGGAGTCTCTCTGGTTGAAAGGGAGCTGAGTACCT 2468
 DB 231 CCAATATGAAACTAGATTAAAGGAGTCTCTCTGGTTGAAAGGGAGCTGAGTACCT 172
 QY 2469 CTGGAACTGGAGGACCTCTGAAAAAGCAAACTGAAACCAAGTGCCTGGGTCACTGTT 2528
 DB 171 CTGGAACTGGAGGACCTCTGAAAAAGCAAACTGAAACCAAGTGCCTGGGTCACTGTT 112
 QY 2529 ACTCTTATAAGACAGTTTAAAGTGAGACCTGAAAAACAATTTGCTTTTACCTTGAATAGAT 2588
 DB 111 ACTCTTATAAGACAGTTTAAAGTGAGACCTGAAAAACAATTTGCTTTTACCTTGAATAGAT 52
 QY 2589 AGGTTTTTATGTTGATATAAGAAATAAACTAATTAACCCCTGAGA 2639

Db 51 AGTTTTTATGTTGATATAAGAAATAAACTAATTAACCTATTAAACCTGAAA 1
 RESULT 11
 BM688117
 LOCUS
 DEFINITION
 UI-E-ClO-abx-a-05-0-UI.r1 UI-E-ClO Homo sapiens cDNA clone
 UI-E-ClO-abx-a-05-0-UI 5', mRNA sequence.
 ACCESSION BM688117
 VERSION BM688117.1 GI:19001375
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 489)
 AUTHORS Bonaudo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.
 Location/Qualifiers
 FEATURES
 source
 1..489
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-ClO-abx-a-05-0-UI"
 /tissue_type="human retina"
 /dev_stage="adult"
 /lab_host="DH10B (Life technologies) (T1 phage resistant)"
 /clone_lib="UI-E-ClO"
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
 modified polylinker; Site 1: EcoR I; Site 2: Not I;
 UI-E-ClO is a cDNA library containing the following
 tissue(s): retina. The library was constructed according
 to Bonaudo, Lennon and Soares, Genome Research, 6:791-806,
 1996. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was ligated to an EcoR I adaptor, digested with Not
 I, and cloned directionally into pT73-Pac vector. The
 oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tag for this library is CCGCG. This library was
 created for the program, Gene Discovery in the Visual
 System, supported by National Eye Institute (NEI)."
 ORIGIN
 Query Match 16.5%; Score 464.8; DB 4; Length 489;
 Best Local Similarity 99.6%; Pred. No. 1.2e-72;
 Matches 466; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2172 TTAAGGATGGAGAAAACCTTCTGCTCTGCGCAAGAAATGGGGAGCTGTGACTGGG 2231
 DB 1 TTAAGGATGGAGAAAACCTTCTGCTCTGCGCAAGAAATGGGGAGCTGTGACTGGG 60
 QY 2232 ATTGTGGGCTGGGCTTGGCTTCTAACTGCTGTGTGACCCAGACAGCCACTTCTCTCCTCC 2291

```

|||||
Db 61 ATTGTGGGCTGGGCTTGGCTTCTAACTGCTGTGTGACCCCAAGACAGCCACTTCTCTCC 120
QY 2292 TAACCTTGGTTATGCTCTTGGCAGCAGAGTGGAGCTAGGAGCAAGAGTGGGATTTGGA 2351
Db 121 TAACCTTGGTTATGCTCTTGGCAGCAGAGTGGAGCTAGGAGCAAGAGTGGGATTTGGA 180
QY 2352 TTATTGTTTTAGATGAGAAATTTATTTTCTTATATATAAACTTATATGTAACCCCA 2411
Db 181 TTATTGTTTTAGATGAGAAATTTATTTTCTTATATATAAACTTATATGTAACCCCA 240
QY 2412 ATATAGAACTAGATTAAGAGGAGCTCTCTGTTGGAAGGAGCTGAGTACCTCTG 2471
Db 241 ATATAGAACTAGATTAAGAGGAGCTCTCTGTTGGAAGGAGCTGAGTACCTCTG 300
QY 2472 GAACCTGGAGCAGCTCTGAAAAAGCAAACTGAAAAACAGTGCCTGGGTCACTGTTACT 2531
Db 301 GAACCTGGAGCAGCTCTGAAAAAGCAAACTGAAAAACAGTGCCTGGGTCACTGTTACT 360
QY 2532 CCTATAGACAGTTTAAAGTGAGACCTGGAAAAACATTTTGCTTTTACCTTGAATAGATAG 2591
Db 361 CCTATAGACAGTTTAAAGTGAGACCTGGAAAAACATTTTGCTTTTACCTTGAATAGATAG 420
QY 2592 TTTTATGTTGGTATATAGAAATTAACCTTAACCTTGAATAGATAG 2639
Db 421 TTTTATGTTGGTATATAGAAATTAACCTTAACCTTGAATAGATAG 468

```

```

RESULT 12
AI424326/c
LOCUS te90c12.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2093974 3'
DEFINITION mRNA sequence.
ACCESSION AI424326
VERSION AI424326.1 GI:4270257
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
cdna Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 578 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 431.
FEATURES
source
location/Qualifiers
1..489
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2093974"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI CGAP Pr28"
/notes="Organ: Prostate; Vector: pT7T3D-Pac (Pharmacia)
with a modified polylinker; Plasmid DNA from the
normalized library NCI_CGAP_Pr22 was prepared, and as
circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization

```

```

reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clones IDs
985608-986759, 110192-110195, and 1217928-1220615).
Subtraction by Bento Soares and M. Fatima Bonaldo. "

```

Query Match 16.4%; Score 461.8; DB 1; Length 489;
Best Local Similarity 99.8%; Pred. No. 4.2e-72;
Matches 463; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

ORIGIN

2175 AAGGATGGAGAAAACCTTCTGCTCTGCGCAAGAATTGGGGAGCTGTTGACTTGGGATT 2234
Db 489 AAGGATGGAGAAAACCTTCTGCTCTGCGCAAGAATTGGGGAGCTGTTGACTTGGGATT 430

2235 GTGGGCTGGGCTTGGCTTCTAACTGCTGTGTGACCAAGACAGCCACTTCTCTCCCTAA 2294
Db 429 GTGGGCTGGGCTTGGCTTCTAACTGCTGTGTGACCAAGACAGCCACTTCTCTCCCTAA 370

2295 CTTGTTTATGCTTGGCAGCAGTGCAGCTAGCGAGCAAGTGGGAGCTGTTGATTGATTA 2354
Db 369 CTTGTTTATGCTTGGCAGCAGTGCAGCTAGCGAGCAAGTGGGAGCTGTTGATTGATTA 310

2355 TTGTGTTTTAGATGTTGGAATTAATTTTGTATATATAAACTTATGTGTAAACCCCAATA 2414
Db 309 TTGTGTTTTAGATGTTGGAATTAATTTTGTATATATAAACTTATGTGTAAACCCCAATA 250

2415 TAGAACTAGATTAAGAGGAGTCTCTGCTGTTGAAAGGGAGCTGAGTACCCCTCTGGAA 2474
Db 249 TAGAACTAGATTAAGAGGAGTCTCTGCTGTTGAAAGGGAGCTGAGTACCCCTCTGGAA 190

2475 CTGGAGCAGCTCTGAAAAAGCAAACTGAAAAACAGTGCCTGGGTCACTGTTACTCT 2534
Db 189 CTGGAGCAGCTCTGAAAAAGCAAACTGAAAAACAGTGCCTGGGTCACTGTTACTCTCT 130

2535 ATAAGCAGCTTTAAAGTGAGACCTGAAAAACATTTTGTCTTACCTTGAATAGATAGTTT 2594
Db 129 ATAAGCAGCTTTAAAGTGAGACCTGAAAAACATTTTGTCTTACCTTGAATAGATAGTTT 70

2595 TTATGTTGTTATATAGAAATTAACCTTAACCTTGAATAGATAG 2639
Db 69 TTATGTTGTTATATAGAAATTAACCTTAACCTTGAATAGATAG 25

```

RESULT 13
BM704154
LOCUS BM704154
DEFINITION UI-B-CK1-afi-p-12-0-UI-rl UI-E-CK1 Homo sapiens cDNA clone
ACCESSION BM704154
VERSION BM704154.1 GI:19017412
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL 97044477
MEDLINE 8889548
PUBMED
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

```



```

RESULT 15
BM686135
LOCUS
DEFINITION BM686135 437 bp mRNA linear EST 27-FEB-2002
UI-E-CKO-aam-b-12-0-UI.r1 UI-E-CKO Homo sapiens cDNA clone
UI-E-CKO-aam-b-12-0-UI 5', mRNA sequence.
ACCESSION BM686135
VERSION
KEYWORDS
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 437)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PubMed 8889548
COMMENT
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.
Location/Qualifiers
1. .437
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-CKO-aam-b-12-0-UI"
/tissue_type="Retina Foveal and Macular"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-CKO"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-CKO is a cDNA library containing the following
tissue(s): Retina Foveal and Macular. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(GT)18 tail. The sequence tag for this library is GTCC.
This library was created for the program, Gene Discovery
in the Visual System, supported by National Eye Institute
(NEI)."
```

FEATURES

```

source
61 CCTCTCTCTCCCAAGGCATCTCACCAGTGGTGTCAACAGGAAATGGCAGCAATAC 120
1982 GTGCAAGGTGTGGTCTTTCTGCTTTGAAAGGGTCAGCTGTACTTTAAAGGGACTGTTTC 2041
121 GCTGCAAGGTGTGGTCTTTCTGCTTTGAAAGGGTCAGCTGTACTTTAAAGGGACTGTTTC 180
2042 AGCTCTGCTGGGTGTGCTGCTGGGACCCCTGCTGCCAACCCACCTCCCTCCCAACAT 2101
181 AGCTCTGCTGGGTGTGCTGCTGGGACCCCTGCTGCCAACCCACCTCCCTCCCAACAT 240
2102 CCTCTCTTTCCATCCATATATCCCTCCAGTATGGACCTTCCCAACTCCCAAGCTGA 2161
241 CTTCTCTTTCCATCCATATATCCCTCCAGTATGGACCTTCCCAACTCCCAAGCTGA 300
2162 ATGTTTCTCTTTAAAGGATGGAGAAATTTCTGTCTGTCTGTGGCAAGAAATGGGGGACT 2221
301 ATGTTTCTCTTTAAAGGATGGAGAAATTTCTGTCTGTCTGTGGCAAGAAATGGGGGACT 360
2222 GTTGACTGGGATTTGGGCTGGGCTTGGCTTCTTAAGTCTGTGTGACCCCAAGACAGCCAC 2281
361 GTTGACTGGGATTTGGGCTGGGCTTGGCTTCTTAAGTCTGTGTGACCCCAAGACAGCCAC 420
2282 TTCTCTCTCTTAACCTT 2298
421 TTCTCTCTCTTAACCTT 437
```

Search completed: June 29, 2005, 05:14:54
Job time : 6078.87 secs

ORIGIN

```

Query Match 15.5%; Score 436; DB 4; Length 437;
Best Local Similarity 99.8%; Pred.No. 1.7e-67;
Matches 436; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1862 GGGCAGCAATAATTTCTATAACTTCCGGGATGCTGAGGGGGGTTTCATGGGAGGACCCCTGG 1921
Db 1 GGGCAGCAATAATTTCTATAACTTCCGGGATGCTGAGGGGGGTTTCATGGGAGGACCCCTGG 60
QY 1922 CCTCTCTCTCCCAAGGCATCTCACCAGTGGTGTCAACAGGAAATGGCAGCAATAC 1981
```

This Page Blank (uspto)